A PROCESS EVALUATION OF THE STATEWIDE IMPLEMENTATION OF THE PYRAMID MODEL FOR SUPPORTING SOCIAL EMOTIONAL COMPETENCE IN INFANTS AND YOUNG CHILDREN IN MASSACHUSETTS

by
Sarah Whitman Geldart

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Education

Baltimore, Maryland
March, 2017

Abstract

Children with disabilities aged three to five in Massachusetts are experiencing poor social emotional outcomes as measured by the Child Outcomes Summary Process. The Massachusetts Department of Elementary and Secondary Education has chosen to implement the Pyramid Model as an evidenced based practice to address these poor outcomes and minimize instances of challenging behavior as part of a State Systemic Improvement Plan. This study examines fidelity of implementation of the Pyramid Model in 18 districts in Massachusetts at three different levels: the state, the district, and the classroom using five distinct criteria for measuring the fidelity of implementation of a program. This study begins with the identification of a problem of practice in the state of Massachusetts, poor social emotional outcomes, and explores the existing literature on this problem. Next, a needs assessment is presented analyzing the prevalence of this problem in the state and the current literature on the Pyramid Model and its use as a practice to address poor social emotional outcomes for children with disabilities. This paper will then outline the key elements of a process evaluation at each level of implementation based on five criteria for measuring fidelity of implementation and describes the various data collection tools used to assess implementation fidelity, including the collection and analysis of information from these tools. These tools are used to assign a rating of high, moderate and low fidelity on each dimension and an overall designation of implementation fidelity for the state. An analysis of Massachusetts' implementation of the Pyramid Model indicates an overall moderate level of implementation fidelity for this initiative. Fidelity of implementation in the state was impacted by several factors including district staff availability, challenges using several of the tools designed to support or analyze implementation, and the structure of the first year of activities. Implications for future research are also discussed including the applicability of the

implementation fidelity matrix created for this study to other statewide implementation fidelity evaluations.

Keywords: fidelity, Pyramid Model, implementation, process evaluation, SSIP

Preface

The research conducted as part of this dissertation is based on my strong interest in supporting states' work in improving outcomes for students with disabilities. In an era of increasing emphasis on results driven accountability it is ever more important for states to be able to define, articulate, and evaluate their improvement activities.

I would like to thank my committee for their guidance and support. Their feedback has helped me to refine my thinking, improve my methodology, and ultimately led to a much better final work. I would also like to thank my colleagues at JHU for their enduring support, ability to always make me laugh when times were stressful, and their ability to motivate me to press on when times were tough. Lastly I would like to extend a special thank you to my husband, David. This process was not without its challenges and regardless of the lost nights together, too many takeout meals to count, and stress of the last three years he has supported me without question. I couldn't have done it without him.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Chapter 1

Introduction

There has been great deal of research indicating that young children with disabilities (CWD) are more likely to demonstrate poor social emotional outcomes in preschool when compared to their nondisabled peers (Anthony, Anthony, Morrel, & Acosta, 2005; Brown & Conroy, 2011; Cannon, Gregory, & Waterstone, 2013; Crnic, Neece, McIntyre, Blacher, & Baker, 2017; Lee, Calkins, & Shin, 2016). The research has shown that improving social emotional skills in early childhood can have long-term benefits for all students, regardless of their disability status (Bulotsky-Shearer & Fantuzzo, 2011). In this context, social emotional outcomes involve how children relate to adults and other children, and follow rules related to groups or interacting with others. It also includes concepts and behaviors such as expressing feelings and emotions, social interactions, and attachment, separation, and autonomy (Early Childhood Technical Assistance Center, 2015).

Every year states are required to report on 17 different indicators of compliance and performance for students with disabilities. These data are collected as part of the State Performance Plan/Annual Performance Report (SPP/APR) process as mandated by the Federal Office of Special Education Programs or OSEP. Included in this report are data on CWD aged three to five in three functional areas: social emotional skills, acquisition of knowledge and skills, and use of appropriate behaviors to meet needs (OSEP, 2006; 2013). Young children in Massachusetts demonstrate poor outcomes in the area of social emotional skills; levels of social emotional skills on statewide child outcomes assessments were lower than the other two outcome areas of acquisition of knowledge and skills and appropriate behaviors to meet needs (Massachusetts Department of Elementary and Secondary Education (MA ESE), 2014a). In Massachusetts, in 2014 only 47.5% of CWD aged three to five

reported in statewide data collection had age expected functioning for the social emotional outcome area by the time they finished preschool. Nationally, 59% of children exit preschool with age expected functioning in this area (Early Childhood Technical Assistance Center, 2015).

Recently, there has been a shift in federal programs serving CWD to emphasize the collection of meaningful student data for both program improvement and accountability purposes (Kasprzak et al., 2011). This emphasis has led states to focus not just on compliance with federal special education regulations, but also on the outcomes of the students they serve, what OSEP terms "Results Driven Accountability" ("Office of Special Education Program's Results Driven Accountability Home Page," 2015). Policymakers are interested in the outcomes achieved for students through participation in programs supported by public funds (Early Childhood Technical Assistance Center, 2015). In addition, President Obama's focus on universal preschool education for all students during his presidency sparked interest in exploring the impact of high quality preschool programs and the benefits and costs of investment in preschool education (The White House, Office of the Press Secretary, 2014). However, these priorities may be shifting under a new presidential administration. At the time of writing a new Education Secretary for the United States Department of Education had just been appointed and the guiding principles of the new administration were still to be determined (Schoen, 2017).

The child outcomes data collection activities, upon which this study is based, were designed in part to meet this demand for individual improvement and accountability from the federal government under the Obama administration. The first challenge that needed to be addressed by the federal government in assessing child outcomes was to identify a universal

measurement tool. These data are collected using the Child Outcomes Summary (COS) Process, a tool developed by the OSEP-funded Early Childhood Outcomes (ECO) Center which has since been renamed the Early Childhood Technical Assistance Center (ECTA). The COS Process is a tool for summarizing a child's functioning across settings and situations, including the results of any relevant assessments. In the process developed by ECTA, the COS Team gathers information about the child's current level of functioning in all of the settings in which they typically spend time. That information is used to assign a rating of 1-7 based on the relationship between the child's current functioning and age-expected functioning in the three outcomes of social emotional skills, acquisition of knowledge and skills, and behavior to meet needs.

In Federal Fiscal Year (FFY) 2012 (school year 2012-2013) OSEP introduced a new performance indicator for all states; Indicator 17: The State Systemic Improvement Plan or SSIP, which states were required to submit initial reporting for on April 1, 2015 (OSEP, 2013). The SSIP is designed to be a comprehensive, yet achievable multi-year plan that is developed in two initial phases (ending in 2015 and 2016) and then implemented in FFY 2015-2018 (OSEP, 2013). The SSIP requires that states assess their current infrastructure capacity and their ability to enhance their infrastructure in order to support the ability of local education agencies (LEAs) to implement, scale up, and sustain, evidenced-based practices that will result in improved outcomes for students with disabilities.

The structure and requirements of the SSIP are based on implementation science, a relatively new field that explores how to promote the systematic utilization of evidence based practices into common practice (ECTA, 2015; Ogden & Fixsen, 2014). The guidance on the development and implementation of the SSIP included explicit references to the stages of

implementation outlined in implementation science and the specific activities states must undertake according to this research in order to ensure the SSIP is implemented successfully (ECTA, 2015). At its core, implementation science is concerned with providing practitioners with the tools and frameworks necessary to translate research-based practices to real world service settings (Ogden & Fixsen, 2014). The foundations of implementation science will be explored in further detail in the literature review.

For the SSIP, MA ESE, in collaboration with statewide stakeholders, chose improving social emotional outcomes for CWD aged three to five as the State Identified Measurable Result (SIMR). MA ESE identified The Pyramid Model for Supporting Social Emotional Competence in Infants and Young Children, also known as the Pyramid Model, as the evidenced based practice to improve social emotional outcomes for young CWD in Massachusetts. There is a strong research base supporting the implementation of this model with early childhood special education students (Blair, 2010; Fox & Smith, 2007; Hemmeter, Snyder, Fox, & Algina, 2016). The Pyramid Model, including training materials on how to implement the model, was developed by two national centers: the Center on the Social and Emotional Foundations for Early Learning (CSEFEL), and the Technical Assistance Center on Social Emotional Interventions (TACSEI; CSEFEL, 2015). The individuals who developed the Pyramid Model on behalf of these centers have subsequently founded a nonprofit organization to promote the high fidelity use of the Pyramid Model, called the Pyramid Model Consortium ("The Pyramid Model Consortium: About Us," 2016). The theory of action for the implementation of the Pyramid Model as part of the SSIP is provided in Figure 1. This study will analyze the implementation of the Pyramid Model statewide in

Massachusetts and through a process analysis identify whether or not this model was implemented with fidelity in the first year of implementation (2015-2016).

This paper will begin with an exploration of why challenging behaviors and poor social emotional development is a concern, and in particular why it is of concern for CWD. It will then move to an examination of the endogenous factors affecting the child including the influence of disability and other child level characteristics. Moving outward from the child, the relevant literature on family and community-level factors influencing these outcomes will be explored. Next, this review will examine the teacher and classroom factors that can impact the social emotional development of young CWD. The literature review will conclude with an analysis of the existing literature on the use of positive behavior supports (PBS) and the Pyramid Model in public school programs. Following the review of the relevant literature, a needs assessment specific to the Massachusetts context is presented. Goals and objectives for this research are identified, followed by an overview of the methodology used to conduct the needs assessment. The methodology section also outlines the author's approach to analyzing implementation fidelity as defined by the five dimensions of fidelity of implementation identified by Dusenbury et al., (2003) and O'Donnell (2008). The results section details the level of implementation fidelity for each of the tools used in this study and each of the five dimensions of fidelity to determine an overall level of implementation fidelity for this initiative. Lastly, the author will present recommendations for future research, a discussion of the results, and an examination of threats to external validity. Appendix A includes a list of acronyms and abbreviations used in this paper for the reader's reference.

Chapter 2

Needs Assessment

In order to provide a more complete review of the nature of social emotional development in early childhood, and particularly for CWD, the following review of the literature is provided. This analysis will offer a theoretical framework in which the reader can understand social emotional outcomes for young children and challenging behaviors. This framework will serve as the basis for the needs assessment to understand the prevalence of poor social emotional outcomes in Massachusetts.

Review of the Literature on Social Emotional Outcomes for Young Children

Social emotional outcomes and challenging behavior. The following review of the literature will focus on social emotional outcomes but will also include a great deal of research on challenging behavior. OSEP and ECTA have defined social emotional outcomes for the COS measurement process to include (a) how children get along with others, (b) how the child interacts with and plays with other children, (c) how children relate to their peers and adults, and (d) how children express emotions and feelings (ECTA, 2015). For children closer to kindergarten age, this outcome area can also include a child's ability to follow rules related to interacting with others in group settings such as child care or the preschool classroom (ECTA, 2015). In this context challenging behaviors can be considered a subset of the global social emotional outcome area. However, there is a great deal of research addressing challenging behaviors specifically and additional research that uses the concepts of challenging behaviors and social emotional challenges interchangeably. In this review, when challenging behaviors are discussed they are referenced as an indicator of poor social emotional skill development and specifically when the research being described focuses on

this subset of social emotional development. Challenging behaviors in children aged three to five include any repeated pattern of behavior(s) that interfere with a child's ability to learn or engage in positive social interactions with their peers and adults. These are behaviors that are not responsive to developmentally appropriate guidance and can include such behaviors as physical and verbal aggression, lengthy tantrums, withdrawal, property destruction, self-injury, and others (CSEFEL, 2015).

Challenging behavior in early childhood. Recently there has been a great deal of interest in the prevalence of challenging behavior for all students, not just those with disabilities. Challenging behavior is a concern due to its rising prevalence among young children and the relationship between social and behavioral competence in young children and their academic achievement (Rimm-Kaufman, Pianta, & Cox, 2000; Schmitt, Pratt, & McClelland, 2014). A large, a national study of 3,593 kindergarten teachers using the National Center for Early Development and Learning's Transition Practices Survey found that almost 21%of teachers indicated that at least half of their class demonstrated problems with social skills upon entering kindergarten (Rimm-Kaufman et al., 2000). Alvarez (2007) studied teacher responses to vignettes depicting student aggression. Responses indicated typical children with behavior problems were rated as more challenging to work with when compared to SWD who do not exhibit comorbid challenging behavior. A more recent study by Hoch et al., (2016) found that children with an identified developmental delay were more likely to demonstrate challenging behavior and that those behaviors would occur more frequently than in their nondisabled peers. Challenging behaviors are of particular concern due to the increased likelihood that these behaviors will lead to students being removed from the classroom.

In 2014, The U.S. Department of Health and Human Services and U.S. Department of Education released a joint policy statement on expulsion and suspension policies in early childhood settings in response to a growing concern about removals in early childhood. Research has shown that children attending preschool programs are being suspended and expelled at high rates (Gilliam, 2005; U.S. Department of Education Office for Civil Rights, 2014).Gilliam, using a random sample of 4,815 classrooms across the country, found that the expulsion rate for behavioral problems in prekindergarten was 6.7 per 1,000, 3.2 times the national rate of expulsion for K-12 students (2.1 per 1,000). These removals are particularly concerning given that school suspension and expulsion practices are associated with negative educational and life outcomes (Lamont et al., 2013; Petras, Masyn, Buckley, Ialongo, & Kellam, 2011). Children who are suspended or expelled at a young age are ten times more likely to be retained in school, experience academic failure, have negative attitudes about school, drop out of high school, and be incarcerated than those who have not (American Psychological Association, 2008; Lamont et al., 2013; Petras et al., 2011). There is also evidence to indicate that removals in early childhood education are associated with suspension or expulsion in later grades (Frey et al., 2015; Mendez, 2003; Poulu, 2015) In addition, removing children from the classroom has the potential to hinder appropriate social-emotional development and to limit access to the cognitively enriching experiences of an early childhood classroom that can lead to academic success and appropriate development later in life (U.S. Department of Health and Human Services, 2014).

Social and emotional skills and students with disabilities. Poor social emotional skills and challenging behaviors are of particular concern for students with disabilities for many reasons. An examination of data from the nationally-representative sample of over

2900 three to five year olds receiving special educations services identified in the Pre-Elementary Education Longitudinal Study (PEELS) found that the prevalence of emotional/behavioral issues was very high in this population (Markowitz et al., 2006). Behavioral problems in structured early childhood learning environments are associated with lower literacy outcomes in kindergarten and first grade (Bulotsky-Shearer & Fantuzzo, 2011). However, a study of 247 preschool children by Scmitt, Pratt, and McClelland (2014) found a positive relationship between directly assessed and teacher-rated behavioral self-regulation and early literacy and math skills.

Social, emotional, and behavioral challenges have also been associated with poor academic achievement in older students with disabilities as well (Cannon, Gregory, & Waterstone, 2013; Education, Families, Interventions, Medicine, & Council, 2009). Nelson, Benner, Lane, and Smith (2004) in a cross-sectional study of a random sample of 155 K-12 students served in public school settings, found that 83% of students with emotional and/or behavioral disorders scored below their non-disabled peers on academic achievement testing in all subjects. Further research of the Special Education Elementary Longitudinal Study (SEELS) reported that students ages 6-21 with social-emotional disabilities are more likely to be excluded from the classroom and excluded multiple times, than students with other disabilities. In addition, these removals appear to lead to additional later removals, as students who were excluded once were much more likely to be excluded again at a later date (Bowman-Perrott et al., 2013). These additional removals further contribute to the likelihood of poor academic and life outcomes described above (Losen, 2014).

There is particular concern for students with disabilities since these students are more likely to demonstrate challenging behaviors than their nondisabled peers. As they age,

students with disabilities are more often removed from the classroom, with data showing that more than 30% of adolescents with disabilities are being suspended or expelled (U.S. Department of Education, 2005). It is also noteworthy that a growing number of children ages birth to three are being reported as demonstrating challenging behaviors in their early intervention programs (Aro, Laakso, Määttä, Tolvanen, & Poikkeus, 2014). Given the prevalence of challenging behaviors, and poor social emotional development in students with disabilities in particular, it is critical to understand factors that contribute to poor social emotional development beginning with the youngest students.

Student level factors. Many factors have been shown to contribute to poor social emotional outcomes for young students. This analysis begins by examining those factors directly related to the student both individually (such as the nature of a disability) and contextually (such as the child's socioeconomic status).

Impact of disability on social emotional development. Early childhood is a critical developmental period for all skills, including social-emotional skills. Developmental neuroscience research has shown that there is rapid growth and development of the areas of the brain responsible for self-regulation and emotion in infancy and early childhood (Nelson & Luciana, 2001). It is important that learning environments at school and in the home foster emotional competence and reduce stress in order to prepare children to develop the self-regulatory and attention necessary for school readiness (Blair, 2002). CWD are likely to exhibit the first signs of poor social emotional development as early as age two or three and it is critical to address these challenges at an early age to prevent later problems (Crnic, Hoffman, Gaze, & Edelbrock, 2004).

Developmental delay is by far the largest disability identified in Massachusetts for children aged three to five (MA ESE, 2014b). Children identified as having a developmental delay are more likely to be identified as having behavioral problems than their nondisabled peers. Researchers have found that CWD are three times as likely to have behaviors that score in the clinical range (Baker et al., 2003; Emerson & Einfeld, 2010). Children with cognitive delay also have a greater prevalence of challenging behaviors at age four than nondisabled peers when controlling for factors of birth circumstances, family, and socio-demographic characteristics such as being born prematurely, having a family that moves frequently, or qualifying as low income (Cheng, Park, Robert, Palta, & Witt, 2014). Children with mild developmental and cognitive delays are also more likely than children without cognitive delays to develop behavior problems and demonstrate poor social emotional outcomes (Crnic et al., 2017; Crnic et al., 2004). Children with developmental delay may not have the cognitive skills necessary to effectively regulate their emotions appropriately and therefore exhibit a wider variety of maladaptive behaviors (Crnic et al., 2004). A 2014 review of the more than 8000 young children that participated in the Early Childhood Longitudinal Study found that children with persistent cognitive delay were more likely to exhibit behavioral problems that did not resolve by the time they reached age five (Cheng, Palta, Kotelchuck, Poehlmann, & Witt, 2014). In addition, behavior problems in young children with cognitive delay are likely to persist into adulthood if they are left unaddressed (Cheng, Park, et al., 2014). These findings are a high priority for Massachusetts since less than 50% of children aged three to five exit preschool special education services with age-expected functioning for social emotional skills.

Communication is the second largest disability category in Massachusetts (MA ESE, 2015). Children with communication disabilities have been shown to demonstrate compromised social skills associated with lower levels of attentional/executive functioning skills (Aro et al., 2014). In addition, children with a speech-language impairment have poorer results on assessments of self-control, assertiveness, sociability, and emotional knowledge understanding (McCabe & Meller, 2004). Children with autism exhibit some of the lowest results on measures of social emotional development among children aged three to five with disabilities. An examination of a nationally representative sample of 2,900 three to five year olds receiving special educations services found that children with autism had significantly lower ratings on the Preschool and Kindergarten Behavior Scales (PKBS-2) – Social Skills scale when compared to other students with disabilities (Ratcliffe, 2014). The communication challenges faced by many young children with autism are especially likely to contribute to the development of problematic behaviors and social emotional challenges in later years (Barber, Saffo, Gilpin, Craft, & Goldstein, 2016; Horner, Carr, Strain, Todd, & Reed, 2002).

Other child level factors. In addition to a child's disability there is a great deal of research exploring how child and family demographic characteristics are related to social emotional outcomes. The PEELS study was one of the most comprehensive studies looking at the variation in the prevalence of disability in young children by different child-level characteristics (Markowitz et al., 2006). This national survey found that almost 70% of preschoolers with disabilities are male and that approximately 67% of CWD aged three to five were white, 22% were Hispanic, and 11% were black/African American. A review of the social emotional development of this population found that males generally exhibited

more problem behaviors than female students and that Black/African American children had higher ratings of challenging behaviors than their White and Hispanic peers (Markowitz et al., 2006). The trend of racial and income gaps for school readiness may be reversing however. A 2016 study of historical data for students born in the late 1990s to the early 2000s by Reardon and Portilla found that while there continue to be differences in school readiness rates by ethnicity, particularly between white and Hispanic students, these gaps have narrowed slightly in recent years. The authors also found that gaps between students identified as low income and those identified as high income have also narrowed in recent years.

Some research has linked race/ethnicity to the likelihood of being identified as having a developmental delay. Researchers found that being Black/African-American or Hispanic in a non-English speaking household is related to higher disability identification rates (Simon, Pastor, Avila, & Blumberg, 2013). What is not clear is if the incidence of disability is actually higher in certain demographic groups or if the determination of eligibility for special education is not culturally sensitive or reflects some other implicit bias.

Similar to the national results, when child outcomes data in Massachusetts is disaggregated by race differences in outcomes become apparent. Black/African American students are more likely to demonstrate lower levels of age expected functioning in the area of social emotional outcomes. This is consistent with research described above and other research that has shown that there are differences in school readiness by student race/ethnicity. While Black/African-American children were more likely to attend early childhood programs than white children, they often attended lower quality programs. In addition, Hispanic children are less likely that white children to attend early childhood

programs (Magnuson & Waldfogel, 2005). A more recent study by National Research Center on Hispanic Children and Families (2016) using data from the 2012 National Survey of Early Care and Education (NSECE) found that Hispanic children aged three to five that are low income were no less likely than children of other ethnicities who also qualified as low income to attend early childhood education programs. However, they did find that this population was less likely to participate in early childhood education programs prior to their third birthday. The lower participation rates in early intervention and/or preschool and the lack of high quality programming for minority students are both possible contributors to poorer outcomes for students, in addition to the other factors described here.

Socioeconomic status and child outcomes. One of the most heavily researched connections between child level characteristics and student outcomes is the relationship between socioeconomic status and social emotional outcomes. There have been conflicting findings as to whether or not children that qualify as low income are over-identified or under-identified as CWD. While some studies have found that poor and minority children nationwide were less likely to receive special education services (Markowitz et al., 2006), others have found that nationally, CWD that also qualify as low income are overrepresented among all CWD (Bethell, Read, Blumberg, & Newacheck, 2007; Simon et al., 2013). What is consistent across the literature is that children who qualify as low income and have a disability have poorer results in measures of health and developmental outcomes when compared to their non-low income disabled peers (Emerson & Hatton, 2007). Recent studies have also emphasized the importance of family engagement over socioeconomic status. Ma, Shen, Krenn, Hu, and Yuan (2016)found positive effect sizes across 46 studies of the

relationship between parental engagement and academic achievement in early childhood education.

Additional research has demonstrated that there may be a link between a child and family's socioeconomic status and the social emotional outcomes of the child. Children from families that qualified as low income were at a higher risk for emotional and behavioral challenges than families that did not qualify as low income (Anthony et al., 2005). Years living in poverty, exposure to household chaos, and inter-parental aggression have been shown to impact young children's ability to recognize and modulate negative emotions (Raver, Blair, & Garrett-Peters, 2014). This is particularly true for CWD. Emerson and Einfeld (2010) found in a large bi-national study of over 35,000 families that children with developmental delays were more likely to exhibit social emotional challenges. In addition, children's exposure to adverse socioeconomic circumstances such as low household income, unemployment, and residing in an area of economic disadvantage were also associated with poorer social emotional development (Magnuson & Waldfogel, 2005).

The relationship between socioeconomic status and social emotional outcomes extends beyond the family to the community as well. An examination of data from the Early Childhood Longitudinal Study of 10,700 children found that children living in economically disadvantaged counties had higher levels of challenging behavior then their peers living in wealthier counties (Cheng et al., 2014). Families that live in economically disadvantaged communities have also been shown to have higher levels of maternal depression and family dysfunction, which are related to negative parenting practices (Kohen, Leventhal, Dahinten, & McIntosh, 2008). Children from families with greater levels of environmental stressors

may also have limited access to educational resources outside of public school programs that could promote positive social skill development (Cheng, Park, et al., 2014).

***Parenting behaviors and social emotional development*.** Parenting behaviors have been shown to impact children's social emotional development. Bennett, Elliott, & Peters (2005) found a relationship between parental stress and child behavior problems. In a longitudinal study of 205 families with 3-year-old children, the authors found that children with developmental delays were more likely to have behavior problems and that there was a relationship between parental stress and behaviors. When stress was high, children were more likely to exhibit problem behaviors. Concomitantly, problem behaviors in children increased with increasing levels of parenting stress as measured by the Family Impact Questionnaire. These results are consistent with reports that parents of CWD often face greater stress than parents of nondisabled students (Green, 2007). Some researchers have hypothesized that the increased demands placed on parents of CWD may contribute to increased negative interactions between parent and child (Crnic et al., 2004). Lastly, parents and guardians of CWD can face substantial hurdles in trying to participate in school and community events. For example, when compared to parents of students without disabilities, families often struggle with issues including multiple care providers, substantial paperwork, fragmented systems of care, stigma, and a lack of inclusive environments (Green, 2007).. Access to high-quality programming and resources for CWD is an important factor in supporting improved outcomes for children and their parents. Whether challenging behaviors and social emotional development are a result of identifiable disability, environmental factors, or both, it is critical that these challenges be addressed in order to maximize children's developmental potential.

Community level factors.

From a socio-ecological perspective the context in which CWD grow up may put them at risk for developing social emotional challenges (Cheng, Park, et al., 2014). Children aged three to five in Massachusetts receive their special education services in what is termed the "mixed delivery system." Statewide, 47.3% of children aged three to five are in full inclusion programs which can include community child care or Head Start programs. Twenty-eight percent (28%) of children are in partial inclusion programs and 15.5% of children receive their special education services in a substantially separate setting. Nine point two percent (9.2%) of children receive their services at home or in some other location (MA ESE, 2014b) . While CWD aged three to five receive their services in a variety of settings, all of these services are provided through or funded by the public school. Nationally, the vast majority of young CWD received speech/language services (93%). Children also commonly received special education services in their preschool (42%), occupational therapy (34%), physical therapy (21%) and individualized tutoring services (19%) (Markowitz et al., 2006). While the public preschool is programmatically and fiscally responsible for the child, special education services can be provided in a number of different settings including the public preschool, Head Start program, family or community-based child care center, or even the home (Individuals With Disabilities Education Act, 20 U.S.C. § 1400, 2004). As such, it is important that the programming in these non-public school locations be aligned with the goals and objectives of the IEP and be high quality. While a state education agency cannot have a direct impact on some of these settings, it is important to understand how environments outside of the public preschool can impact social emotional development.

Child care quality and social emotional development. Only three studies to date, and all in Australia, have explored child care educators knowledge and comfort with addressing social emotional concerns, including mental health issues. One study explored baseline data around mental health literacy in early childhood programs (Farrell & Travers, 2005). The authors found that educators lacked the knowledge and capacity to promote the mental health of children attending childcare. In addition, while educators' awareness of risk and protective factors for mental health issues increased immediately after training, this knowledge was not sustained over a 12 month period. A second study also examined educator's mental health literacy and found that they had challenges identifying the causes and early signs of mental health problems in children (Williamson, Davis, Priest, & Harrison, 2011). Davis et al. (2014) used a survey to assess 24 child care providers. She found that while educators expressed confidence in their knowledge and ability to support children with social emotional challenges, measures of classroom quality indicated opportunities for improvement in supporting children's social emotional development. In addition, there was no significant correlation between quality of care and educators' self-reported knowledge. Massachusetts recently developed and released social emotional early learning standards for public and private early childhood programs. The roll out of these standards, along with high quality professional development to support their sustained implementation, could potentially support increased knowledge of appropriate teaching practices to support improved social emotional development in young children.

Teacher level factors. As a state education agency, MA ESE can have a large impact on teacher preparation and professional development of licensed educators through statewide initiatives, regulatory changes, funding, and professional development opportunities (MA

ESE, 2014a). Through educator preparation programs, teacher licensure, and ongoing professional development and technical assistance, MA ESE can provide targeted support to licensed public educators serving CWD aged three to five. The studies outlined below address the importance of teacher preparation, professional development and high quality teacher/child interactions in supporting the social emotional development of young children.

Teacher preparation. Most early childhood special education teachers have completed higher education coursework as part of their teacher preparation program. Nationally, more than half (55%) of CWD aged three to five have had a teacher with a graduate degree and 38% had a teacher with a bachelor's degree (Markowitz et al., 2006). However, the programs offered by institutes of higher education may not focus specifically on young CWD. A survey of faculty members from higher education programs in nine states analyzed how faculty prepare early childhood educators to address young children's social emotional development (Hemmeter, Santos, & Ostrosky, 2008). Faculty from both two and four year institutions reported their students had only emerging skills related to children's social emotional development. Respondents indicated a need for training materials on how to address the needs of young children with challenging behaviors that could be incorporated into existing curriculum. These results are consistent with studies by Early et. al. (2007) and Mashburn et al. (2008) showing that teacher degree attainment and participation in higher education coursework was not directly correlated with performance in the classroom supporting early childhood outcomes.

Similarly, a meta-analysis of seven studies of preschool programs did not find a relationship between teachers' educational attainment or school major with the outcomes of classroom quality or children's academic outcomes (Early et al., 2007). This indicates that

there may not be a correlation between a teacher's pre-service preparation program and the children's outcomes in these teacher's classrooms. The authors noted that one concern is that many teacher preparation programs often do not include an early childhood component as described above. In addition, teachers might not be provided adequate support at the school and classroom level to implement the effective practices they have learned. Early et al. found that teacher quality is not simply measured by degree attainment and that additional in-service professional development may be necessary to help teachers provide a high-quality early education experience.

The prevalence of challenging behaviors can vary significantly across classrooms. Anthony et al. (2005) offers several possible explanations for this phenomenon. First, differences in ratings may reflect differences in teacher training or experience. Second, it may also be that teachers do not understand how to appropriately measure social emotional functioning or they may apply different standards for what age appropriate functioning looks like. It has been shown in preschool children that the variation in problem behaviors across classrooms emerges 2-3 months after children begin school. This may indicate that it is more likely a difference in teachers' skills and abilities to deal with challenging behavior that is causing this difference in the prevalence of challenging behaviors rather than differences in the children in each classroom (Anthony et al., 2005). These results are consistent with a recent survey of a sample of over 10,000 educators working with children under the age of five. The study found that approximately 20% of these individuals received training on supporting children's social emotional development in the last year (National Survey of Early Care and Education, 2013). Given the lack of pre-service preparation in early childhood

special education programs described above, the need for additional training may be much higher than the current rate of 20%.

The evaluation of social emotional outcomes at the state level in Massachusetts is heavily dependent upon a teachers' ability to accurately assess students' current levels of functioning (ECTA, 2015). There is limited research on a teacher's ability to interpret curriculum-based measurements and use that data to develop observable and measurable objectives in a students' Individual Education Plan (IEP). One small study of three teachers at a private special education school suggested that teachers may have difficulty in taking assessment data and using it to develop appropriate, meaningful IEP goals (Codding, Skowron, & Pace, 2005).

Another common concern about teachers is the lack of training in child development and the ability to distinguish between behaviors that are developmentally appropriate and those that are not (Qi & Kaiser, 2003). Knowledge of developmentally appropriate behaviors is context dependent for both teachers and parents. Researchers have found differing ratings of social competence for children with communication delays on norm-referenced assessments between parents and teachers and depending on the environment in which a child is assessed (McCabe & Meller, 2004). When working with children that have a disability, it is particularly important that teachers have training in recognizing which behaviors are a manifestation of a child's disability and to understand how to make reasonable accommodations and modifications for those children so that they are not impermissibly suspended or expelled as a result of behaviors related to disabilities (34 C.F.R. §§104.4, 104.38, 28 C.F.R. § 35.130(b)(1), 7).

High quality professional development. Challenging behaviors in early childhood classrooms are often a core concern for educators. A small study of five preschool special education teachers in North Carolina found that the most pressing need for support among this group was additional training in how to manage the challenging behaviors of their students (Gebbie, Ceglowski, Taylor, & Miels, 2012). This need for additional support to manage challenging behaviors in students with disabilities is present regardless of the number of years of teaching experience a teacher has (Gebbie et al., 2012). A review of state practices in implementing the COS process as a measurement for early childhood outcomes found that educators need professional development and technical assistance in a variety of topics, including: "child development, functional outcomes, authentic assessment, data collection techniques (family interviews and observation), how to interpret and use data, and the implementation of promising practices related to service delivery models, services in natural environments, and early childhood transition" (Kasprzak et al., 2011, p. 161). One of the greatest challenges for states implementing outcomes measurement in early childhood special education as a requirement under the Individuals with Disabilities Education Act is the integration of this measurement practice into existing state systems and supporting ongoing professional development for educators (Kasprzak et al., 2011).

Data collection and evaluation is an important aspect of special education. However, teachers and service providers can struggle with collecting data consistently and using that data to inform their instruction (Sandall, Schwartz, & Lacroix, 2004). Teachers often reported dissatisfaction with their data collection methods and their ability to use data to inform instruction and monitor children's progress (Sandall et al., 2004). A review of assessment practices related to child outcomes measurement in Pennsylvania provides

valuable information on how the process is implemented and how educators feel about this data collection. Campbell & Anketell (2007) used focus groups and surveys to gather information from approximately 500 individuals across the state. The results indicated that few assessments were administered using the recommended best practices and professional standards for early childhood assessment, which may lead to less reliable assessment results. This study also revealed a hesitation among practitioners to use a single standardized measure. Practitioners were concerned about the granularity of measures in detecting minor changes in student performance with less regard for whether or not the assessment used was designed or validated for use with students with disabilities. The authors argued that it is important for educators to utilize measures that have been validated for use with students with disabilities and that have the ability to detect small changes in performance to account for the sometimes slow progress this population makes. This study also points to potential concerns about the use of the COS process as a way to assimilate information about a child's current functioning if the assessments used by educators are not used accurately or are inappropriate for the population being assessed (Campbell & Anketell, 2007).

There is only one peer-reviewed study of the implementation of the COS Process by COS teams. Greenwood, Walker, Hornbeck, Hebbeler, and Spiker (2007) found that a variety of educators collaborated to determine the outcomes ratings for individual students, and teams included service providers as well as teachers. Due to the variety of services this population receives, alignment across providers may be critical to ensure the consistency of service delivery.

Because of the prevalence of communication disorders in children aged three to five, speech-language pathologists (SLPs) have an important role to play in the development of

social emotional skills in this population. SLPs are responsible for a great deal of authentic and functional assessment and can be a significant source of information for the COS process. SLPs are well trained in high-quality assessment and can contribute to the valid assessment of child outcomes, including social emotional skills (Hebbeler & Rooney, 2009).

Teaching assistants are also an important part of many early childhood classrooms that serve students with disabilities. A recent review of the existing literature revealed that there is limited research on the role of support staff, including how they can be used to support the inclusion of students with disabilities with their non-disabled peers (Saddler, 2014). There is little existing research on the impact these staff members can have on student outcomes. Therefore, it is difficult to quantify the role that these individuals can play in supporting the social emotional outcomes of young CWD.

Paraprofessionals also play a significant role in early childhood special education. However, it is likely that these individuals need additional pre-service or in-service training to adequately support CWD in an inclusive setting. In part because of their lack of training, some paraprofessionals have challenges demonstrating a consistent ability to tailor instruction to the individual needs of their students, including instruction in social emotional skills (Schepis, Ownbey, Parsons, & Reid, 2000).

***Quality teacher/child interactions*.** One measure of classroom and teacher quality is the nature of teacher-child interactions. Higher quality emotional interactions have been associated with improved social competence and fewer challenging behaviors for young children (Mashburn et al., 2008). In addition, the extent to which teachers interact positively with children and provide coherent instruction, including scaffolding and informative feedback is associated with positive academic outcomes, including social emotional skills, by

the end of kindergarten (Domínguez, Vitiello, Fuccillo, Greenfield, & Bulotsky-Shearer, 2011). Classroom process quality, including high emotional support by teachers, has also been shown to buffer the detrimental effects of problem behavior in early childhood classrooms (Domínguez et al., 2011).

Markowitz et al. (2006), in a national survey of over 2,900 teachers, found that a majority of teachers believe that they are engaging in positive relationships with their students and supporting social interactions. For teachers of CWD aged three to five, 89% reported that they provided structured play and task situations. In addition, 77% reported that they encouraged CWD to initiate and maintain interactions with their non-disabled peers However, while these teachers are supporting peer-to-peer interactions, they might not be focusing on teacher-child interactions. A study of 240 randomly selected publicly funded preschool programs in six states found that prekindergarten teachers on average did not engage in high-quality interactions with their students or offer clear, content-rich instruction (Burchinal et al., 2008). This is particularly important given that the same study found that high-quality teacher-child interactions are associated with improved social outcomes in kindergarten. In addition, problems in teacher-child interactions in preschool have also been linked to lower reading and language achievement in first grade (Bulotsky-Shearer & Fantuzzo, 2011). While teachers may feel comfortable supporting engagement between peers, they may need additional professional development and support to engage in high-quality, supportive interactions with their students, which has been shown to be a protective factor for social emotional development.

Classroom level factors. Classroom climate and characteristics can also impact a child's social emotional development. In particular, the prevalence of challenging behaviors

or the perception of their prevalence can impact the outcomes of students in the classroom: "the effect of a classroom climate in which more children are displaying difficult behaviors and less social competence, or at least, in which teachers perceive more children with these problems, would be expected to have significant effects on the academic and social progress of children in the classroom" (Anthony et al., 2005, p. 37). A child's ability to maximize the value they get out of their early childhood education is dependent in part upon their ability to effectively engage with their peers, teachers, and the tasks of the classroom (Vitiello, Booren, Downer, & Williford, 2012). Children who engage in challenging behaviors during structured learning time exhibit poorer cognitive and social emotional outcomes, possibly due to the lost opportunity to learn (Domínguez et al., 2011). If there are challenges in the dynamics of the classroom or it is not designed to support this population, children may not receive the maximum benefit of their education.

Classroom engagement can also play an important role in the social emotional development of young children. Vitiello et. al. (2012) found that children were more engaged in chosen activities and more engaged with teachers when participating in teacher-driven activities. While this study did not examine the relationship between engagement and problem behaviors it does suggest the importance of structuring a classroom to facilitate learning through increased engagement.

Discussion of the literature on poor social emotional outcomes. There is a great deal of literature on the prevalence of poor social emotional outcomes for CWD aged three to five. What is evident from a review of this literature it that there is likely no single cause for these poor outcomes and therefore no single solution. While a child's disability may make it more likely for that child to struggle to develop age-appropriate social emotional skills, other

factors such as race and socioeconomic status are related to social emotional development independent of disability. Family level factors such as inter-parental conflict and parental stress have also been shown to impact student outcomes.

MA ESE is a state education agency, it can serve as a strong leverage point to improve the poor social emotional outcomes for CWD aged three to five in the public school system (Cohen & Hill, 2008). Many teacher preparation programs do not have adequate coursework in early childhood special education or in how to improve social emotional outcomes for all children. In addition, research on current teachers has shown that they need additional support to accurately collect data to assess student's performance, to interpret that data, and to individualize instruction based on the data analyses. Many teachers also need support in working with children with challenging behaviors to help them remain in the classroom. Teachers should also be encouraged to engage in positive relationships with their students. Finally, classroom engagement and appropriate classroom quality standards may support improved social emotional outcomes in this population but the research on this area is somewhat mixed.

While there is a great deal of research addressing the prevalence of poor social emotional outcomes among students with disabilities and how certain personal and contextual factors can contribute to poor outcomes for non-disabled students, there is very little information on how the personal and contextual factors beyond disability can impact the social emotional development of young CWD. In addition, there is not a significant amount of research on early childhood settings outside of public preschool and Head Start programs. Many children in Massachusetts attend private or family-based childcare and it is important to understand how these environments may also impact a child's development.

Review of the Literature on Implementation Science

The key improvement strategy for the SSIP selected by MA ESE was the implementation of the Pyramid Model. According to OSEP, the structure of the SSIP, including the required components were modeled after the principles of implementation science (ECTA Center, 2015). Implementation science is primarily focused on understanding how to take evidenced-based practices and implement or scale up their use in practical settings. The need for implementation science is in part due to studies that found current dissemination and diffusion strategies for evidenced-based practices have resulted in the limited use of these programs (Ogden & Fixsen, 2014) Rather than focus on outcomes for individual participants, implementation science examines the components of a process necessary to support sustainable and scalable use of an evidence-based intervention (Ogden & Fixen).What follows is a brief overview of implementation science as it relates to the implementation of the SSIP, including support for the implementation of an evidenced-based practice statewide.

The foundations of implementation science lie in a seminal study by the National Implementation Research Network in 2005 (Bertram, Blase, & Fixsen, 2015). This study laid out a framework of key intervention components that can serve as the foundation for the selection, clarification, and implementation of an intervention or evidenced-based practice. These components include: a strong definition of the model; an identification of the theory supporting the activities and elements of the model; a theory of change; an identification of the target population's characteristics; and a rationale for why alternative models were not chosen (Bertram et al., 2015).

This study also established several stages of implementation. According to Bertram et al., (2015), implementation is a process than unfolds over the course of two to four years and begins with an exploration stage in which an organization assesses their needs, examines the components of an intervention, assesses the fit between an intervention (including its target population and required resources) and their own context, and considers what would be needed to implement a given intervention. The second stage is the installation stage in which resources are utilized and structural supports are put in place that are necessary precursors to implement the new practice (Bertram et al., 2015). The third stage in implementation science is the initial implementation stage. At this stage the evidence-based program is implemented and challenges often arise that require modifications to the implementation plan based on stakeholder and participant feedback. The final stage is full implementation in which practitioners are regularly using the evidenced-based practice with fidelity (Bertram et al., 2015).

Implementation science has also utilizes the term "implementation drivers" which are those factors that establish the capacity for implementation of an intervention effectively and with high fidelity (Bertram et al., 2015). There are three types of implementation drivers: competency drivers, organization drivers, and leadership drivers (Fixsen, Blase, Metz, & Dyke, 2013). Competency drivers support the competence of practitioners and include staff selection, coaching, professional development, and performance assessment (Bertram et al., 2015). Organization drivers are those that help to create an environment that provides administrative, fiscal, political, and procedural environments that ensure competency drivers are able to be utilized effectively and monitored for effectiveness(Bertram et al., 2015). Leadership drivers include both technical and adaptive leadership to respond to challenges in

which there is a great deal of agreement about a strategy to address the challenges and those in which there is less certainty and agreement respectively. Regardless of the type of challenges, leadership drivers help to identify the challenges, the possible solutions, and articulate a way forward (Bertram et al., 2015). Each of these three drivers are interdependent and must be consistently monitored to support high quality of implementation and fidelity to the model as designed in order to ensure the successful implementation of an initiative (Bertram et al., 2015).

Scaling Up and Sustaining Evidenced-Based Practices in Districts.

The research that establishes a practice as evidenced-based is often based on small scale studies or implementation in a specific context (Klingner, Boardman, & McMaster, 2013). Initiatives like the SSIP require states to take an evidenced-based practice and scale it up to multiple environments, each with their own unique characteristics and contexts. In order for a district to be successful in implementing a new initiative the initiative must meet local needs and be responsive to contextual factors within the district(Klingner et al., 2013). This includes ensuring that any new practices can be easily integrated into existing work and that teachers implementing these practices understand the importance of specific practices and how they may lead to improvements (Klingner et al., 2013)

Researchers have identified a number of factors that support systems change and the adoption of and sustainability of evidenced-based practices in districts. These include maximizing the fit between a district's needs and the practice, ensuring that the practice is a priority among stakeholders, ensuring fidelity of implementation, integrating the practice with daily school or district operations, and finally, embedding a cycle of continuous improvement through effective data collection and analysis (Klingner et al., 2013).There are

a number of challenges associated with taking an initiative to scale across multiple districts including dealing with a more heterogeneous student population, service providers and teachers of varying skill and knowledge levels, differing levels of resources and attention to an initiative, and varying levels of support (Ogden & Fixsen, 2014). District personnel may also present a challenge with scaling up an initiative as they do not always prioritize or understand the importance of utilizing research-based programs (Klingner et al., 2013).

Professional development and scaling up

Odom (2009) coined the term "enlightened professional development" to refer to the transformational practices required to scale up and sustain evidenced-based practices across environments as part of an implementation science framework. Enlightened professional development incorporates a strong focus on the context in which practices will be adopted, incorporates collaborative processes to adopt the new practices, utilizes coaching and consultation, and provides opportunities for educators to discuss issues that arise during implementation (Odom, 2009). This model also emphasizes the importance of long-term relationships between those seeking to support the implementation of a practice such as researchers and the individuals implementing a practice. Lastly, unlike traditional professional development models, the enlightened professional development model requires trainers with expertise in the content and the ways in which the trainer can facilitate active engagement and community building to sustain the use of these practices (Odom, 2009).

Implementation Science and Statewide Implementation.

Fixsen, Blase, Metz, & Dyke (2013) developed a framework for the statewide implementation of evidenced-based programs based on the implementation science model. The authors argue that too much emphasis has been placed on establishing evidenced-based

practice while little focus has been given to establishing the statewide infrastructure necessary to support these practices at scale. The statewide implementation framework developed by Fixsen et al., (2013) places equal emphasis on the intervention itself and the development of the systems and infrastructure to support implementation of the intervention. They argue that in addition to the implementation team, there must be external supports for system change, policies to enable effective practice, and practice-policy communication to truly create system change. One of the greatest threats to the sustainability of any program is the dilution or drift that occurs over time as services are delivered in a way that is inconsistent with the model as originally designed (Ogden & Fixsen, 2014). This challenge is particularly relevant for statewide implementation as resources to support sustained engagement with participants may be limited.

The stages of implementation described previously can unfold at a variety of levels including the federal, state, and local, each with their own unique context, drivers, and challenges. The fidelity of implementation of a given program can be impacted by the alignment or misalignment of the priorities and or systems at each of these levels (Bertram et al., 2015). In order to ensure successful, effective implementation it is important for administrators at the state level to analyze and address those factors that may support or constrain implementation fidelity and the targeted outcomes of an initiative (Bertram et al., 2015).

Need Assessment to Understand Poor Social Emotional Outcomes in Massachusetts

Research Question: Examining Social Emotional Skills in Massachusetts.

The above review of the literature illustrates the prevalence of challenging behaviors and poor social emotional outcomes for CWD aged three to five. In addition to presenting challenges in the immediate classroom context, these behaviors can contribute to poorer academic achievement over time (Rimm-Kaufman et al., 2000). To better support the social emotional needs of children served in Massachusetts, this needs assessment seeks to examine how social emotional outcomes for students with disabilities aged three to five vary across time and contexts. Specifically the following research questions will be addressed:

1.0 How have the COS performance values in Massachusetts for social and emotional skills (Outcome A) changed across time (2006-2013) in comparison to state targeted values and national results?

1.1 What has been the percent of children who increased their rate of growth by program exit (Statement 1) across time in comparison to state targeted values and national results?

1.2 What has been the percent of children functioning within age expectations by program exit (Statement 2) across time in comparison to state targeted values and national results?

2.0 How do 2012/13 COS performance values in Massachusetts for social emotional skills (Outcome A) vary by child demographic and service delivery contexts?

2.1 Do COS performance values of child progress outcomes defined by Summary Statements 1 and 2 for social emotional skills vary by Massachusetts state accountability and assistance levels for school performance?

2.2 Do COS performance values of child progress outcomes defined by Summary Statements 1 and 2 for social emotional skills vary by child race?

2.3 Do COS performance values of child progress outcomes defined by Summary Statements 1 and 2 for social emotional skills vary by child ELL status?

2.4 Do COS performance values of child progress outcomes defined by Summary Statements 1 and 2 for social emotional skills vary by child disability?

2.5 Do COS performance values of child progress outcomes defined by Summary Statements 1 and 2 for social emotional skills vary by child need/service intensity?

2.6 Do COS performance values of child progress outcomes defined by Summary Statements 1 and 2 for social emotional skills vary by child placement?

Method

Participants

Children three to five years of age who received ECSE services in Massachusetts under Part B and who participated in entry and exit data collection in specific years were participants.

Setting and cohort model.

The data presented below were collected using a representative sample of children in Massachusetts. MA ESE collects the child outcomes data using a cohort model. Each of the state's approximately 400 districts are assigned to one of four cohorts and each cohort has a unique data reporting requirement every year for the State Performance Plan / Annual Performance Report submitted to OSEP (MA ESE, 2014). Child outcomes data are collected for one cohort each year beginning in August and ending in December. Entry data is collected and reported to MA ESE for any child aged three to five who is new to special education services during that time. Exit data for the sample of students is reported to MA ESE by districts each June until all of the students for whom entry data was collected have

exited the program (MA ESE, 2014). OSEP has approved this sampling methodology as a valid and reliable way of collecting child outcomes data (MA ESE, 2014). The comparisons of the Massachusetts data to national data are based on the publically available data through ECTA.

Variables.

Indicator 7 data is reported using five progress categories and two summary statements for each of the three outcome areas based on the ratings at entry, the ratings at exit, and the answer to the progress question for each outcome at exit. The five progress categories for federal reporting are defined as follows: (a) the percent of children who did not improve functioning; (b) percent of children who improved functioning but not sufficient to move nearer to functioning comparable to same-aged peers; (c) percent of children who improved functioning to a level nearer to same-aged peers but did not reach it; (d) percent of children who improved functioning to reach a level comparable to same-aged peers; and (e) percent of children who maintained functioning at a level comparable to same-aged peers (OSEP, 2013). These five progress categories are intended to capture every potential type of progress or regression a child can make in a program. Reporting on five different progress categories for the three outcomes leads to a great deal of information but can also make it challenging to compare longitudinally and across programs. Therefore, OSEP collapses the five progress categories into two summary statements for each outcome area:

- **Summary Statement 1:** Of those preschool children who entered or exited the preschool program below age expectations in each Outcome, the percent who substantially increased their rate of growth by the time they turned 6 years of age or

exited the program ((Progress Categories C + D) / (Progress Categories A + B + C + D).

- **Summary Statement 2:** The percent of preschool children who were functioning within age expectations in each Outcome by the time they turned 6 years of age or exited the program ((Progress Categories D + E) / (Progress Categories A + B + C + D + E). (Federal Office of Special Education Programs, 2013).

The results presented below utilize the summary statement values for ease of interpretation.

Procedure.

Analyses of the current status of social emotional outcomes in Massachusetts were conducted using existing data collected by this researcher as part of their regular employment. Analyses were limited due to restrictions around the use of data containing individually identifiable student information that cannot be taken from the MA ESE building due to data privacy regulations. Therefore all charts and tables presented in this needs assessment have been repurposed from other publically available presentations and were made available through a public records request of the researchers' own work. The analyses presented in the results section below include a sample of 1840-1867 students depending on the information available for each student in the Massachusetts Student Information Management System (SIMS). All disaggregated data presented below was collected and reported in Federal Fiscal Years 2012 and 2013 (MA ESE, 2014). Longitudinal data was collected using the data publically available in the SPP/APR.

Design and analysis.

This study utilized a historical retrospective design to examine the COS performance changes over time of CWD aged three to five. Descriptive and Central Tendency statistics were used to graph data to address the research questions.

Results

Research question 1.0.

Research question 1.0 asked how the COS performance values in Massachusetts for social emotional skills (Outcome A) have changed across time in comparison to state targeted values and national statistics. This research question was investigated using two subquestions.

Research question 1.1 addressed what percentage of children across time increased their rate of growth by program exit (Summary Statement 1) in comparison to state targeted values and national results. This research question was analyzed by examining the percent of children in each year who were identified as having substantially increased their rate of social skills growth by exit (see Figure 3). While it is clear that growth has occurred for Summary Statement 1 values from FFY2006 to FFY2013, there has been some fluctuation in this growth. In addition, while there was initial growth in Summary Statement 2 values from FFY2006 to FFY2011, the most recent two years have seen a decline in this data. Figure 4 shows that gains have been made in the percentage of children who demonstrate substantial growth by exit (Summary Statement 1) in social emotional skills in Massachusetts for Outcome A, in comparison to national statistics, but the gains have not met the targets set by Massachusetts.

Research question 1.2 asked what has been the percent of children who demonstrate age-expected social skills functioning by program exit (Summary Statement 2) across time,

in comparison to state targeted values and national statistics? This question was addressed by examining the percent of children in each year who were identified with social skills functioning within age expectations by exit (see Figure 3). While it is clear that growth in social skills functioning occurred from 2006 to 2011, from 2011 there has been approximately a 10 percentage point decline in age-expected social skills functioning at exit. Figure 5 presents this flat and slightly declining trend more clearly across time with social skills performance consistently below the national average and well below targets set by Massachusetts.

Research Question 1.0 asked if the COS performance values in Massachusetts for social emotional skills (Outcome A) changed across time in comparison to state targeted values and national statistics. While the number of children in Massachusetts demonstrating improved social emotional functioning at exit across time has increased, surpassing national averages, the improved functioning has not been sufficient for children to meet age expectations in this outcome area by the time children exited. It was clearly demonstrated that age appropriate social skills functioning for children was below national averages and was not on track to meet state targets.

Research question 2.0.

Research Question 2.0 asked how the 2012/13 performance values in Massachusetts for social emotional skills (Outcome A) vary by child demographic and service delivery contexts. This research question was addressed using 6 subquestions.

Research question 2.1 asked if COS performance values of child progress defined by Summary Statements 1 and 2 for social emotional skills vary by Massachusetts state accountability and assistance levels for school performance. This question was first

addressed by examining the average entry and exit ratings of children in each accountability and assistance level (1-5, with the lowest performing schools in Level 5) as presented in Figure 6. Students in Level 1 districts had the highest average entry (*M*=4.3) and exit ratings (*M*=5.5) on the COS Process scale of 1-7 for the social emotional outcome but the least average change in functioning from entry to exit (*M*=1.18). In contrast, students in Level 4 districts, the lowest performing districts in this sample, had the lowest entry rating (*M*=3.5) but the highest change from entry to exit (*M*=1.65) of the four accountability levels. One limitation of the disaggregation by District Accountability and Assistance Level was the unequal distribution of students across the four levels. In this sample 68 children in 15 different LEAs are represented in the Level 1 category, 1086 children in 137 LEAs were in the Level 2 category, 500 children in 41 LEAs are in the Level 3 category, and 213 children in 9 LEAs are in the Level 4 category. This disproportionate weighting of students in Level 2 and the small number of districts in Level 4 may impact the applicability of these results.

Research Question 2.1 was also addressed by examining the percent of children in each state accountability and assistance level who demonstrated greater than expected growth (Summary Statement 1) and the percent who exited at age expected functioning (Summary Statement 2) in the social emotional outcome area. Figure 7 shows that children attending Level 1 and Level 4 schools were more likely to demonstrate greater than expected growth in the social emotional outcome area. In addition, children in Level 3 schools were the least likely to exit at age expectations in this outcome area.

Research question 2.2 asked whether COS performance values of child progress defined by Summary Statements 1 and 2 for social emotional skills vary by child race. This question was addressed by examining the percent of children disaggregated by race who

were identified in Summary Statements 1 or 2 (see Figure 8). The percentage of students making greater than expected growth in social emotional outcomes ranged from a low of 75% for multiracial students to a high of 86% of students for Hispanic/Latino and black/African American students. However, multiracial students had the second highest percentage of students exiting at age expectations at 49%. White students had the greatest percentage of students exiting at age expectations at 54%. Black/ African American students were the least likely to exit at age expectations with 42% meeting this criteria. An analysis of the representativeness of the sample indicates that the percentage of students by race in the sample is consistent with the statewide population of three to five year old students with disabilities.

Research question 2.3 asked whether COS performance values of child progress defined by Summary Statements 1 and 2 for social emotional skills vary by child ELL status. A review of the ELL identification information led the researcher to determine that this data was not valid and reliable for students aged three to five as districts with known large ELL populations reported no ELL students in their data. This is likely attributable to ELL status criteria that can be challenging to apply for children only attending school part-time or for special education services alone (T. Valentine, personal communication, April 14, 2015). Figure 9 presents the results for native and non-native English speakers, a proxy for ELL status. While there is a sight difference in Summary Statement 2 values there is a larger difference in the Summary Statement 1 values with 90% of non-native English speakers and 83% of native English speakers showing greater than expected growth. Research question 2.4 asked if COS performance values of child progress defined by Summary Statements 1 and 2 for social emotional skills vary by child disability. This question was addressed using data

from children participating in exit data collection in 2012 and 2013 which was disaggregated by their identified disability in the state student information system (see Figure 10). In Massachusetts the three largest disability categories for students ages three to five are developmental delay, communication, and autism respectively (Massachusetts Department of Elementary and Secondary Education, 2015). Figure 10 shows that only 32% of students with autism demonstrated age-expected functioning by the time they exited preschool. In contrast, 48% of students with a developmental delay and 66% of students with a communication disability met this benchmark. There was less variation in the percentage of students who made greater than expected growth.

Research question 2.5 asked do COS performance values of child progress defined by Summary Statements 1 and 2 for social emotional skills vary by child need/service intensity? This question was addressed using data from children who participated in exit data collection in 2012 and 2013. The data was disaggregated by a student's identified level of need. The level of need variable is pulled from a student's IEP in Massachusetts and is based on the number of hours of service a student receives a week and the IEP Team's professional judgment about this variable (T. Valentine, personal communication, April 14, 2015). This information is presented in Figure 9. There is a strong association between level of need and the percentage of students exiting at age expectations (Summary Statement 2). The higher a student's identified level of need, the smaller the percentage of children who exited at age expected functioning in the social emotional outcome area. A small amount of variation in the Summary Statement 1 values across level of need can be observed. There is a seven percentage point difference in the percentage of children making greater than expected growth from the students with the lowest identified level of need to the highest.

Research question 2.6 asked whether COS performance values of child progress defined by Summary Statements 1 and 2 for social emotional skills vary by child placement. This question was addressed by examining the percent of children in each placement category who were identified as either having substantially increased their rate of growth (Summary Statement 1) or exited at age expectations (Summary Statement 2) in the social emotional outcome area (see Figure 12). Massachusetts defines placement for three to five year old children in the following ways. Full inclusion includes children that attend any inclusive program in which more than 50% of the students do not have IEPs and the child receives a majority of their special education services in that inclusive setting. This can include the public preschool program, Head Start, community childcare, and other settings (Massachusetts Department of Elementary and Secondary Education, 2014). Partial inclusion means that a child attends an inclusive setting but receives a majority of their special education services outside of the inclusive setting. Substantially separate indicates that a child attends a public or private program or classroom specifically designed for students with disabilities and the home/service provider category indicates that the child does not attend any program but receives their special education services either at home or in the service provider's office. Children in partial and full inclusion programs were more likely to exit at age expectations in this outcome area. Children in substantially separate settings were the least likely to have age expected skills in this area by the time they exit preschool but an analysis of the correlation between placement and level of need indicates a strong correlation between placement and level of need with a correlation coefficient of .79.

In summary, Research Question 2.0 asked how the 2012/13 performance values in Massachusetts for social emotional skills (Outcome A) vary by child demographic and

service delivery contexts. The analysis presented previously demonstrates that there is variation in performance on Summary Statement 1 and 2 values when the data is disaggregated by different subgroups, particularly in the percentage of children exiting at age expectations (Summary Statement 2). The differences in the percentage of children exiting at age expectations were greatest when the data was disaggregated by disability, level of need, and placement. Children with autism and children with a high level of need were the least likely to exit with age expected functioning in the social emotional outcome area. Children receiving services in the home, with the lowest level of need, and/or with a communication disability were the most likely to have age expected functioning in the social emotional outcome area by the time they exited preschool.

Discussion of the needs assessment.

Based on the analyses presented in this needs assessment, there is clear evidence that CWD aged three to five are exhibiting poor social emotional outcomes, both when compared to other outcome areas (acquisition of knowledge and skills and behavior to meet needs), and when compared to the national data. A review of the disaggregated data by several key factors identified in the literature shows trends that are consistent with the existing literature. Children in high-quality programs tend to have better outcomes than those in districts that rank lower on the state's Accountability and Assistance Level system. Children with autism and higher levels of need overall tend to perform worse than students with other disabilities or less need and students spending at least part of their day with non-disabled peers are more likely to have age expected social emotional skills.

Based on this identified need in Massachusetts, a literature review was conducted to understand existing practices to address social emotional challenges for young children with

disabilities. What follows is a review of the literature on the Pyramid Model for Supporting Social Emotional Competence in Infants and Young Children as a possible evidenced-based practice to support improved social emotional outcomes for young children in Massachusetts.

Literature Review of the Pyramid Model

The Pyramid Model was developed by the Center on the Social and Emotional Foundations for Early Learning (CSEFEL) and is a tiered conceptual framework of evidence-based practices for supporting social emotional development in young children, including the creation of positive school climate (CSEFEL, 2015). The foundation of the Pyramid Model relies upon the systems and policies in schools and programs necessary to ensure a workforce can adopt evidenced-based practices. The first tier includes universal supports for all children including nurturing and responsive relationships and high quality environments. The second tier is prevention, including targeted social emotional strategies to prevent problems. The third tier represents individualized intensive interventions for specific children who demonstrate the most challenging behaviors (CSEFEL, 2015). The Pyramid Model is interchangeably referred to as Preschool PBS, PBS, and the Pyramid Model and therefore this literature review explores relevant works utilizing any of these three terms in early childhood programs.

While the Pyramid Model and PBS more generally are often seen as a general education initiative, the origins of this tiered intervention framework can be found in special education. The development of what we now know as PBS began in the late 1980s with a grant from the U.S. Department of Education National Institute on Disability and Rehabilitation Research (NIDRR) to a consortium of universities to develop non-aversive behavior management strategies for individuals with disabilities (Johnston, Foxx, Jacobson,

Green, & Mulick, 2006). Further funding by OSEP, the creation of a national Positive Behavior Interventions and Supports (PBIS) center, and explicit efforts to refine this model for use in early childhood programs has led to a strong presence from PBS in special education (Johnston et al., 2006). However, the emphasis on preventative strategies and whole-school behavior management in PBS has led to its adoption by educators of all children (Bradshaw, Reinke, Brown, Bevans, & Leaf, 2008).

Challenges to implementation fidelity and the existing literature.

PBS is a tertiary intervention framework that incorporates a number of evidenced-based practices to support children with challenging behaviors (Fox & Smith, 2007) and has been widely used and studied in school-age populations (Marshall, Brown, Conroy, & Knopf, 2011). However, because early childhood special education services can be provided across a variety of settings in the mixed-delivery system including Head Start, private child care, and home based care, the PBS model has been implemented somewhat differently for young children (Marshall et al, 2011.). While there has been a great deal of research exploring the implementation of this model in individual preschool-age children, there is far more limited research on the expansion of this model program-wide and teacher's ability to implement a constellation of evidenced-based practices with fidelity (Muscott, Pomerleau, & Szczesiul, 2009).

Several recent studies point to the challenges of implementing PBS program-wide. Benedict, Horner, & Squires (2007) examined the implementation of PBS in 15 early childhood settings. They found that teacher's use of PBS strategies were limited and that even with additional coaching in PBS practices, challenging behaviors in classrooms were not substantially diminished, possibly due to low incidences of these behaviors overall prior

to intervention. At an even broader scale, Muscott et al. (2009) examined the implementation of PBS in 47 public and private early education programs across New Hampshire. Similar to Benedict et al. (2007), they found that teachers were able to successfully implement some but not all components of the model and struggled the most to use features designed to address the most challenging behaviors. Ongoing coaching of teachers implementing PBS can be used to support the use of this model by educators. Fox, Hemmeter, Snyder, Binder, & Clarke (2011) studied the impact of ongoing coaching for three teachers implementing PBS. While they were able to identify a functional relationship between coaching and implementation of PBS practices, they still experienced challenges ensuring implementation with fidelity.

Each of these studies points to a common challenge in program-wide implementation of PBS. While there is substantial research on the implementation of PBS for school-age children across relatively similar school structures (Marshall et al., 2011), the varying nature of the types and structures of early childhood classrooms makes it challenging to conduct studies that are applicable across programs. In addition, the fact that PBS is a program consisting of a constellation of evidenced-based practices can make it a challenge to implement with fidelity, even with additional coaching support. Lastly, these inconsistencies in implementation can make it challenging to take the program to scale across multiple programs (Benedict et al., 2007; Marshall et al., 2011; Muscott et al., 2009). Given these challenges, it is important to examine the evidence base supporting the implementation of the Pyramid Model in preschool programs serving children with and without disabilities.

Literature Review Method

In order to understand the evidence base supporting the implementation of the Pyramid Model in preschool classrooms a comprehensive literature search was conducted. First, the researcher completed a comprehensive search of four electronic databases: Education Resources Information Center (ERIC), Google Scholar, and Education Full Text, and PsychInfo. Various combinations of the following search terms were used: PBS, Positive Behavior Supports, Positive Behavior Interventions and Supports, PBIS, Pyramid Model, CSEFEL, early childhood, and preschool. Second, the researcher examined the reference lists of these articles for additional studies. Only articles from peer-reviewed journals published in the last fifteen years were included in each of these searches. After removing duplicates and articles not related to both PBS and early childhood education, the author was left with 47 studies. The researcher conducted a systematic evaluation of each article to ascertain its relevance to this study. This included looking at each study's methodology, participants, and results. The author then identified several criteria to further narrow the field to those studies most relevant for analyzing the evidence base supporting the use of PBS and/or the Pyramid Model in public preschool classrooms. In order to be included in the analyses below a study must specifically state that public preschool children participated in the intervention. Studies that included children in preschool in addition to other grades were included. Each study must also explicitly mention the use of either PBS or the Pyramid Model as the intervention. Refining the search based on these criteria left the author with eleven articles.

The author used the quality standards developed by Gersten et al. (2005) to analyze the quality of the remaining eleven studies in order to determine the overall evidence base supporting the use of the Pyramid Model in public preschool classrooms and with students with disabilities in particular. These standards were selected because of their applicability to

special education research and because a majority of the studies identified through the comprehensive search described above involved group experimental or quasi-experimental studies.

Gersten et al. (2005) identified ten essential quality indicators and eight desirable quality indicators for group experimental and quasi-experimental research articles and reports. The ten essential quality indicators are grouped into four distinct domains: how the authors describe participants, implementation of the intervention, outcomes measures, and the data analyses used in each study. According to Gersten et al., authors should describe participants in a study in such a way that it is easy for the reader to determine whether or not participants demonstrated the disability(ies) noted. In addition, the reader should be able to evaluate the appropriateness of the procedures used to identify comparable samples and conditions in addition to sufficient information necessary to understand the characteristics of the interventionists. High quality studies according to Gersten et al. clearly describe the nature of an intervention and how it was implemented. They also include a clear description of how fidelity was measured. The outcomes measures for high quality studies for these authors include the use of multiple measures at appropriate times. Gersten et al. identify the need for appropriate data analysis techniques and effect size calculations in their studies. The eight desirable quality indicators included elements that while not essential, are likely to contribute to a higher quality study.

In order for a study to be considered high quality according to Gersten et al. (2005) it would need to meet all but one of the identified essential quality indicators and at least four of the desirable quality indicators. These same criteria were applied to the eleven studies under examination here. The content and methodology of each of the eleven studies

identified by the author were analyzed for the presence of each essential and desirable quality indicator and coded using the standards established by Gersten et al. in their work. If a quality indicator was met by a study then the article was assigned a score of one on that indicator. If it was not met then it was assigned a zero. The results of these analyses, including each of the quality indicators identified by Gersten et al., are described later.

The mean domain scores were calculated for each of the four domains similar to the methodology used by Nagro and Cornelius (2013). The domain scores were calculated by totaling the number of sub-domains that met the high quality criteria (totaling the points) and dividing by the total possible number of points in a domain across the eleven studies. For example, the quality indicator for describing participants domain has three distinct indicators that a study could possibly meet (Gersten et al., 2005). Across eleven studies, there are a total of 33 possible points that could be assigned in this domain. However, of the eleven articles examined, only 23 of these subdomains were met across the identified studies. Therefore, for the eleven studies under examination here, 66.7% of the criteria on this domain were met. The results of these analyses can be found in Figure 2 below. This analysis by domain helped to identify areas of strength and areas of weakness in the existing research as outlined below.

Literature Review Results

Research findings. As described above, all of the identified studies addressed the use of PBS or the Pyramid Model in early childhood classrooms. An overview of the content of each study, including study design, participants, the nature of the intervention and the results are presented in Table 1. Decreases in challenging behavior among young children receiving positive behavioral supports were seen across several of the studies (K.-S. C. Blair, 2010; Duda, Dunlap, Fox, Lentini, & Clarke, 2004; Hall et al., 2007; Stoiber & Gettinger, 2011;

Voorhees, Walker, Snell, & Smith, 2013). Many of the identified studies examined the impact of professional development and coaching on the fidelity of implementation of PBS and the Pyramid Model (Benedict et al., 2007; K.-S. C. Blair, 2010; Carter & Van Norman, 2010; Carter, Van Norman, & Tredwell, 2011; Duda et al., 2004; Fox et al., 2011; Howard S. Muscott, Mann, & LeBrun, 2008; Howard Steven Muscott et al., 2009). Overall these studies found that implementation fidelity and use of the practices associated with PBS was improved with additional professional development and coaching. Lastly, only one study examined the impact of training in PBS on families. Hall et al. (2007) found that through intensive support in PBS strategies families reported lower levels of overall stress in addition to a reduction in challenging behaviors on the part of their children.

Research quality. Table 2 shows the numeric scores for the eleven identified studies on each of the essential and desirable quality indicators as identified by Gersten et al. (2005). As noted above, if a study met a quality indicator it is assigned a numeric rating of one, if it did not it is assigned a zero. In order for a study to be identified as high quality according to the criteria established by Gersten et al. it must meet all but one essential quality indicators and at least four desirable quality indicators. Four of the eleven studies examined here met this criteria: Blair, Fox, and Lentini, (2010); Carter and Van Norman (2010); Stoiber and Gettinger (2011); and Voorhees, Walker, Snell, & Smith (2013). A majority of the remaining studies did not meet the criteria for high quality because they had more than one essential quality indicator unmet.

Discussion. Gersten et al. (2005) proposed that there be at least four acceptable quality studies or two high quality studies supporting a practice in order for it to be considered evidenced based. As described above, four of the eleven studies analyzed meet

the criteria for high quality studies thereby supporting the notion that the use of PBS in preschool classrooms is an evidenced based practice. For those studies who did not meet the acceptable or high quality criteria, most studies did not adequately describe or even utilize a control condition for comparison as shown in Table 2. This also contributed to the likelihood that a study would not meet the desirable quality indicator on the documentation of instruction in comparison condition.

It is notable that across these eleven studies a variety of techniques were used to analyze the effects of PBS in preschool classrooms. Studies ranged in scope from working with individual student teacher pairs to an examination of statewide implementation of PBS. These widely different experimental designs might have led to the unequal distribution of high quality and desirable indicators being met across the studies. Figure 2 provides a visual display of the grand mean scores for the eleven studies in each of the four essential quality indicator domains established by Gersten et al. (2005). The studies examined here were most likely to meet the criteria for high quality in their measurement of outcomes, the grand mean score on this domain was 86.4%. Results on the other three domains ranged from 66.7% to 72.7%.

While the mixed results on the high quality criteria and the desirable quality indicators may be due to the quality of the research conducted, it is also likely that it reflects the quality of reporting in the articles. In some cases conducting the analyses to determine whether or not a criterion was met was challenging due to the vague language or inconsistent reporting included in a given article. The use of PBS in early childhood settings is relatively recent, implemented within the last fifteen years (Fox & Smith, 2007). Therefore, it is likely that as this field expands and develops a stronger literature base there will be additional studies that

meet the high quality research criteria described here. The variety in the design of the studies also proves challenging for researchers trying to apply this literature base to particular contexts. Replicating research across different environments with similar methodology and participants will build a stronger literature base that will have additional relevance for early childhood educators. Expanding use of the Pyramid Model in early childhood education will produce more studies focusing on implementation in this specific context. However, given the explicit attention given to clearly articulating best practices and recommendations for practitioners by the authors of these studies, any individual or program seeking to implement PBS has access to a wealth of knowledge and experience through the studies examined here.

Conclusions

Chapter two established the prevalence of the problem addressed by this study, poor social emotional outcomes for CWD ages three to five, and reviewed the extant literature on poor social emotional outcomes and the use of the Pyramid Model as an evidence-based practice to address poor social emotional outcomes. The review of the literature on young CWD indicated that this population was more likely to demonstrate challenging behaviors and poor social emotional outcomes than their nondisabled peers. In addition, the research has established the importance of classroom practices and teacher preparation in supporting improved outcomes, and that those improved outcomes can contribute to greater academic success over time. A needs assessment on the prevalence of this problem in Massachusetts indicated that CWD aged three to five in the state are exhibiting poor social emotional outcomes both when compared to other outcome areas (acquisition of knowledge and skills and behavior to meet needs) and when compared to the national data. Lastly, a review of eleven studies of the Pyramid Model indicates establishes that it is an evidence-based

practices for supporting the social emotional development of young CWD according to the criteria established by Gersten et al. (2005). Based on the analyses presented in this chapter, the following evaluation question was established for this study.

Evaluation Question

To what degree are MA ESE, the eighteen participating Pyramid Model districts, and the classrooms within those districts, implementing the Pyramid Model with fidelity in each of the five dimensions of fidelity of implementation identified by Dusenbury et al., (2003) and O'Donnell (2008) in the first year of statewide implementation (2015-2016)?

Chapter 3

Method of Program Evaluation of the Implementation of the Pyramid Model in Massachusetts

MA ESE has chosen to implement the Pyramid Model as part of the SSIP to improve social emotional outcomes for young children with disabilities aged three to five. This study will use a program evaluation framework to examine the ways in which Massachusetts has approached the implementation of the Pyramid Model statewide.

Participants and Setting

This program evaluation will use extant data from an ongoing initiative sponsored by MA ESE. The implementation of the Pyramid Model is part of a statewide initiative called the State Systemic Improvement Plan (SSIP; MA ESE, 2014). The SSIP is a comprehensive, multi-year plan to improve social emotional outcomes for young children aged three to five with disabilities in Massachusetts that is submitted to OSEP as part of the SPP/APR reporting process described previously. MA ESE began work on the SSIP in 2014 and the work is ongoing. The theory of action developed by MA ESE for this implementation is provided in Figure 1. Consistent with the theory of action, MA ESE's implementation plan consists of activities at three levels: the state, the district/program, and the classroom. At the state level, MA ESE has contracted with the Pyramid Model Consortium to advise the state on implementation in the 18 participating districts, provide trainings consistent with the model, and offer guidance and support for expansion into additional districts. MA ESE has also contracted with eight individuals with expertise in the model to act as external coaches for each district. Districts are paired with an external coach who can advise district-based leadership teams and classroom coaches on the implementation of the model. Each external

coach is responsible for coaching between one and four districts. In addition, MA ESE is working with the Pyramid Model Consortium to develop the ability of these external coaches to train in this model.

Each participating district has identified a district-based leadership team to oversee the local implementation of the Pyramid Model. These teams participate in statewide training activities, coaching from the external coaches, and strategic planning sessions with individuals from the Pyramid Model Consortium. The leadership teams are responsible for supporting a scalable and sustainable implementation of this model in the district. Lastly, at the classroom level, educators are participating in statewide trainings in the Pyramid Model practices and receiving ongoing coaching support from classroom coaches. Classroom coaches are identified by the district leadership team, and receive training in Pyramid Model practices and practice-based coaching.

Materials

When the Pyramid Model was first developed, a number of key measures of implementation fidelity were developed to support implementation of this model consistent with its original design (CSEFEL, 2015). In addition, MA ESE has the benefit of working directly with the creators of this model through its contract with the Pyramid Model Consortium to receive direct assistance and guidance on implementation in the state. Outlined below are the data collection tools that will be used to measure implementation fidelity at the state, district and classroom level, many of which were developed by the creators of the Pyramid Model. Table 3 presents an overview of each data collection tool used in this study including a description of the tool, the data source, the frequency at which data is collected, and the criteria for high, moderate, and low fidelity on each. Each of these

tools can be used to establish fidelity of implementation at the state, district, and/or classroom level under the five criteria established by Dusenbury et al., (2003) and O'Donnell (2008) described later. What follows is a brief overview of each tool. The next section will articulate how each of these measures can inform an understanding of fidelity of implementation in each of the five criteria.

SSIP phase I and phase II reports. As described in the introduction, MA ESE is required to submit an annual report to OSEP as part of the SPP/APR reporting process (OSEP, 2013). Included in this report is a detailed description of the activities MA ESE has undertaken or plans to implement as part of the SSIP. The Phase I report includes information on how MA ESE identified poor social emotional outcomes for young children with disabilities as the focus of its SSIP (MA ESE, 2015). The Phase II report includes an implementation plan, evaluation plan, and an overview of how MA ESE intends to support districts and classrooms implementing the Pyramid Model (MA ESE, 2016). The principal investigator of this study is the primary author of the Phase I and II reports. OSEP provides states with feedback on the alignment of the submitted reports, including plans for implementation, with OSEP requirements and whether or not all necessary components are present (OSEP, 2013). High fidelity of implementation on this tool is indicated by full alignment with OSEP requirements and an indication that all required components are present.

Training materials. A key component of fidelity at all levels of implementation is adherence to the Pyramid Model, as originally designed, including during statewide training activities for external coaches, district leadership teams, classroom coaches, and educators implementing the model in their classrooms. Data for this measure will be collected by

analyzing training materials including PowerPoints, handouts, and evaluations from participants as trainings occur to ensure consistency with the materials developed by the Pyramid Model Consortium. Engagement in trainings, including what, if any, information needs to be clarified or included in future training activities will be evaluated using training evaluations provided by participants. High fidelity of implementation is strong alignment with the standards and materials established by the Pyramid Model Consortium and a strong indication of understanding of the content by participants at the trainings. Low fidelity of implementation at the state level would be characterized by a lack of adherence to the materials, scope, and sequence established by the national Pyramid Model Consortium.

Interviews. MA ESE's contract with the Pyramid Model Consortium includes ongoing consultation with the developers of this model to ensure that implementation in Massachusetts is consistent with the design of the model and meets the needs of Massachusetts educators. Interviews with the developers as they offer statewide trainings and consult on the development of the Phase I and Phase II SSIP reports will inform analyses of implementation fidelity. High implementation fidelity would be indicated by their feedback as developers of this model. A copy of the interview questions is included in appendix B.

Training participant data. In order for districts and educators to implement the Pyramid Model with fidelity they will need to receive training in the model (Fox & Smith, 2007). The dose and reach of training received by the 18 participating districts, their educators, and coaches will be measured by participation in statewide training activities such as the District Leadership Team Academies, Practices Trainings, and Coaches Trainings. Participation in statewide training activities will be measured by training registration and attendance. In order for the Pyramid Model to be implemented program wide, all educators in

each program must be trained in the model. The evaluation of the training data will demonstrate the number of educators in each of the 18 districts that have participated in high quality training activities and whether or not those individuals represent the intended audience for a particular activity.

District benchmarks of quality. The Pyramid Model Consortium has developed a Benchmarks of Quality document for programs and districts implementing the Pyramid Model (L. Fox, personal communication, March 15, 2015). This self-assessment identifies 47 benchmarks of quality in nine critical element categories on which programs can assess their implementation on a scale of not in place, partially in place, or in place. As O'Donnell (2008) notes, fidelity of implementation to a program wide reform model, such as the Pyramid Model, is more challenging than fidelity of implementation for an individual educator implementing a curriculum in the classroom. Program wide implementation requires system change and is more likely when there is an intervention manual, such as the Pyramid Model Benchmarks of Quality, that outlines the critical features of the model (O'Donnell, 2008a). Participating districts complete the Benchmarks of Quality when they attend initial Pyramid Model Leadership Team Academies. Thereafter they complete the tool at least three times a year: at the beginning, middle and end of the school year. External coaches support the district leadership team in the completion of this measure and can provide additional guidance on the meaning of each element and what it means for each element to be *in place*, the criteria for fidelity on the tool (L. Fox, personal communication, March 15, 2015). This will support more accurate self-assessment by each participating district. For this study, high fidelity of implementation on this tool is demonstrated by a district indicating that at least 75%of the critical features of the Pyramid Model are *in place*

in their district. A copy of the District Benchmarks of Quality form is included in Appendix B.

Teaching Pyramid Observation Tool (TPOT). Hemmeter, Fox, and Snyder (2013) have developed and published a classroom implementation fidelity tool for the Pyramid Model called the Teaching Pyramid Observation Tool for Preschool Classrooms (TPOT). According to the authors, the TPOT utilizes a two hour classroom observation and subsequent interview with the teacher to score implementation of key Pyramid Model practices in 14 areas, identify the presence of *red flags* during instruction or in the classroom, and understand an educator's response to challenging behavior. A pilot version of the TPOT was evaluated in an efficacy trial with 50 preschool classrooms, and the results suggested that the TPOT is a reliable measure for teacher's implementation of the Pyramid Model (Snyder, Hemmeter, Fox, Bishop, & Miller, 2013). A score of 80% or higher on the TPOT indicates implementation fidelity, according to the authors of this measure (Hemmeter, Fox, & Snyder, 2013).

External coach contact record. A critical component of the Pyramid Model is ongoing practice-based coaching of educators and programs implementing the model (Fox et al., 2011). MA ESE has developed an extensive External Coach Contact Record for external coaches under contract with MA ESE. External coaches complete this online form each time they have a substantive contact with a district. This measure captures information including the mode of contact (e.g., site visit, phone call, email), the type of support provided, the number of individuals trained, and the status of the district's current Pyramid Model implementation. This tool provides ongoing information about what kinds of support districts need, and how many classrooms and teachers are trained, and have adopted the Pyramid

Model. The data collected is coded and analyzed for themes. These aggregate analyses will inform decisions about training topics, training frequency, and support for external coaches. A copy of the External Coach Contact Record is included in Appendix B.

Evaluation Framework

Evaluating five dimensions of implementation fidelity. Dusenbury, Brannigan, Falco, and Hansen (2003) and O'Donnell (2008) present five dimensions of fidelity that must be measured in order to understand whether a program is being implemented as intended including (a) adherence; (b) duration or dose; (c) quality of delivery; (d) participant responsiveness; (e) program differentiation. The authors note that it is unclear whether or not all five types of fidelity must be present in order for a program to be successful. However, a careful examination of how the data tools described previously can inform measurement of fidelity of implementation at the state, district, and classroom level will provide a comprehensive perspective on the use of the Pyramid Model in Massachusetts at all levels of implementation. Table 4 presents the six data collection tools, a description of each tool, the source of data for the tool, the frequency at which that tool is utilized, the individual responsible for collecting the data using the tool, and definitions of high, moderate, and low implementation fidelity for use of the tool. Table 5 presents the five criteria for fidelity, a list of tools that inform each dimension, and the requisite element(s) from each tool that informs that dimension of fidelity. Overall, as shown in Table 5, there are 21 intersections between the six tools utilized in this study and the five dimensions of fidelity being assessed.

Adherence. Adherence is the extent to which the essential components of an intervention are implemented as intended (Dusenbury et al., 2003; O'Donnell, 2008). There are a number of data tools that will inform an analysis of adherence. At the state level,

fidelity of implementation is the adherence to the model's design or the extent to which state level training and coaching activities align with the MA ESE theory of action, SSIP, and the materials and guidance provided by the Pyramid Model Consortium (Dusenbury et al., 2003). The SSIP Phase I and II reports will provide qualitative and quantitative data on the implementation plan and its execution, including the theory of action for this intervention, and information on how activities were implemented. The training materials for statewide training activities and interviews with the national Pyramid Model Consortium consultants will inform an analysis of whether or not the evidenced-based materials developed by the creators of this model are being used as intended, and if the trainings are delivered in a manner consistent with the design of the intervention. At the district level the training materials will support an analysis of whether or not participants are receiving the appropriate trainings as planned. Training participant data will elucidate whether or not the intended audience is participating in the training. The District Benchmarks of Quality will provide information on the alignment between the local level activities that support the implementation of this model and the model's identified critical features. The External Coach Contact Record provides information on district and classroom-level implementation of the key components of the Pyramid Model to assess adherence. Lastly, the TPOT will be used as a tool to measure whether or not teachers are using activities and methods consistent with the design of the Pyramid Model.

Dose and Duration. Dusenbury et al., (2003) and O'Donnell (2008) define the measurement of dose and duration of an intervention as the amount of program content received by participants. This can include information on the number, frequency, and duration of training activities. At the state level, the SSIP report will provide information on

the statewide activities that take place as part of Massachusetts' implementation of the Pyramid Model. Similar to the analysis of adherence, the training materials will provide information on the dose and duration of each training activity to analyze fidelity at the state and district level. The External Coach Contact Record includes information on the number, type, and length of training activities that occur at the district level, and the number and roles of the individuals attending trainings. At the classroom level, data on the number of TPOTs completed for each teacher will demonstrate the extent to which classroom-based coaches are engaging in coaching activities with implementing educators.

Quality of Delivery. The quality of delivery of an intervention is the third critical element to assess for implementation fidelity. Quality of delivery is the extent to which "the implementer delivers the program using the techniques, processes, or methods prescribed" (O'Donnell, 2008, p. 34). The contract with the Pyramid Model Consortium is a critical aspect of the evaluation of quality of the delivery of the Pyramid Model in Massachusetts. This contract includes direct training to educators statewide from the creators of this model which ensures the quality of delivery for these activities. In addition, the contract stipulates that the national trainers provide virtual and face-to-face support for MA ESE's External Coaches to develop their capacity to provide high-quality trainings and support. Through ongoing interviews with the Pyramid Model Consortium the principal researcher will be able to assess the quality of the delivery at the state level. In addition, a review of the training materials will demonstrate whether or not trainers are using the critical materials and activities of the Pyramid Model in their trainings, and if the participants indicate training is of high quality. The district Benchmarks of Quality can be analyzed as an indicator of whether or not the critical processes and techniques of the Pyramid Model are being utilized as part of

a district's implementation. The External Coach Contact Record can inform whether District Leadership Teams and classroom teachers are using the intended methods to implement the Pyramid Model practices. For example, are Leadership Teams using a collaborative process to facilitate their monthly meetings? Are teachers making meaningful connections between their posted visual schedule and classroom activities throughout the day? As an implementation fidelity assessment, the TPOT provides information on the quality of delivery of Pyramid Model practices in the classroom. This assessment examines how the key practices of the model are used in the classroom and whether or not practices that can limit the effectiveness of the model, what are termed *red flags* in the TPOT, are present in the classroom.

Participant Responsiveness. Participant engagement in an intervention is a critical element for success (O'Donnell, 2008). Participant responsiveness refers to the degree to which individuals are involved in and utilizing the components and activities of an intervention (Dusenbury et al., 2003; O'Donnell, 2008). Training evaluations provide information about how individuals perceive the statewide training activities and whether or not they are likely to attend similar events in the future. Similarly, training participant data can indicate whether or not individuals remain engaged in the initiative through their ongoing attendance at statewide training events. The completion of the District Benchmarks of Quality on a regular basis is an indication of the engagement of District Leadership Teams and the results on this measure can indicate whether or not districts continue to make progress on their implementation of the Pyramid Model. The External Coach Contact Record will provide documentation of district and teacher engagement through a neutral observer.

External coaches will be asked to submit information on school and classroom level activities that support the use of Pyramid Model practices using the External Coach Contact Record.

Program Differentiation. An examination of the fidelity of implementation through the lens of program differentiation is intended to establish whether or not the critical features of a program that set it apart from a control condition are present or absent as part of an implementation. Evaluating program differentiation at the state level for the implementation of the Pyramid Model as part of the SSIP in Massachusetts will not be possible. Massachusetts was a partner state in the initial development of the Pyramid Model by CSEFEL and TACSEI over 12 years ago, and this participation has led to the Pyramid Model practices being utilized in some form in a significant portion of the early childhood classrooms in the state (T. Williams Valentine, personal communication, March 1, 2015). The 18 districts participating in this initiative are receiving substantial supports including free statewide training, coaching, and some additional funding (MA ESE, 2014). However, the state has not identified a control or comparison condition to use in the evaluation of this initiative. The District Benchmarks of Quality can be used to identify which critical elements are in place in a program and which are absent but as the study is designed, does not allow for comparison between implementing and non-implementing districts. A pre- and post-completion of the Benchmarks of Quality does allow for an analysis of changes in implementation over time. Dusenbury et al. (2003) emphasize the importance of identifying which elements should be excluded as part of program differentiation. The TPOT calls these elements *red flags* and provides classroom-based coaches with the opportunity to coach educators about the elimination of these activities as they occur in implementing classrooms.

Procedure

Extant data collected by the author of this study according to the following timeline will be examined to identify the fidelity of implementation according to the five dimensions of fidelity described previously.

Design

A key aspect of the SSIP is the ongoing monitoring of implementation of this program for continuous improvement so that MA ESE can modify implementation as necessary to ensure it is implemented with high fidelity (OSEP, 2013; Resnick et al., 2005 as cited in O'Donnell, 2008). As an efficacy study, this study will examine whether or not the Pyramid Model was implemented as intended as detailed in the SSIP implementation plan and theory of action. For this initiative, implementation fidelity will be measured at the state, district, and local levels as previously described in the evaluation framework using the data collected as described in Table 1.

Analysis

For this study the author will use documentation from the development of the SSIP and the implementation of the Pyramid Model in Massachusetts from January 2015 through September 2016. This timeframe covers the initial drafting of the Phase I SSIP report, a planning document outlining the implementation plan for the Pyramid Model in Massachusetts through the end of the summer following the 2015-2016 school year in which all 18 participating districts began implementation of the model in their programs. As the lead manager on this statewide initiative the author has access to all source documents for this project and the vendors under contract with MA ESE. The Phase I and II SSIP reports have already been submitted to OSEP and MA ESE is awaiting written feedback on the

Phase II report from OSEP which will be used to inform the analysis of implementation fidelity at the state level. All extant training materials from the following trainings will be cross referenced with published materials from CSEFEL and the Pyramid Model consortium and analyzed for any deviations from the evidenced based content: three statewide leadership team meetings, three 4-day Pyramid Model Practices trainings, two TPOT trainings, two Prevent, Teach, Reinforce Young Children trainings, and two series of five, one hour coaches calls. Written notes from calls within this timeframe with the Pyramid Model consortium on the implementation of this model in Massachusetts in addition to supplementary interviews if necessary will be analyzed to assess whether or not the developers of this model believe it to be implemented with fidelity in Massachusetts. Extant training participant data will be analyzed using frequency counts and descriptive statistics to identify the audience for each training. These participants will then be compared to the recommended participants for each training. The 18 participating districts completed the Benchmarks of Quality at the beginning of their participation in this program, in the middle of the 2015-2016 school years, and in June 2016. These results will be analyzed in aggregate and by district to determine the extent of changes at the program level in the implementation of this model. MA ESE is collecting TPOT scores in aggregate by district to understand implementation fidelity at the classroom level. While a limited number of TPOTs were completed in the timeframe of the study, this information will be used to inform an analysis of the initial fidelity of implementation at the classroom level. Lastly, all external coach contact records for this time frame will be analyzed to document both quantitative reports of implementation activities and the qualitative accounts of the quality of that implementation.

An overall determination of high, medium, or low fidelity of implementation for this initiative will be established using the criteria presented in Table 7. Table 7 presents a matrix of the five dimensions of fidelity as established by O'Donnell (2008) and each of the tools identified previously for assessing implementation fidelity of the Pyramid Model as shown in Table 5. For each tool and relevant dimension of fidelity, the author has established criteria for high, moderate, and low fidelity and the points assigned to each of these ratings (3, 2, and 1 respectively). While each data collection tool informs more than one dimension of fidelity, not all tools inform every dimension of fidelity. For example, interviews with the developer of this model can inform an understanding of the adherence of this initiative to the model as designed but cannot speak to participant responsiveness in this initiative. It should be noted that this initiative did not utilize a control group and therefore none of the data collection tools inform program differentiation or the differences between participating districts and those not implementing the Pyramid Model.

This matrix will be used to assign an overall fidelity of implementation score for this initiative based on a total of the number of points the initiative earns each time there is an intersection between a data collection tool and a dimension of fidelity. For example, the SSIP reports (data collection tool) inform an analysis of adherence (dimension of fidelity) to the model. Full alignment with the Phase I Implementation Plan for the evidenced-based practice and OSEP requirements would earn the initiative 3 points for high fidelity while misalignment with the Phase I Implementation Plan, the OSEP Requirements, and/or more than four articulated components of the evidence-based practice missing from the implementation would lead to a score of one for low fidelity of implementation on adherence to the model. Overall there are 21 intersections in this matrix between a data collection tool

and a dimension of fidelity and therefore 21 opportunities for a district to score a maximum of three points for high fidelity or 63 points overall. In order to achieve an overall composite rating of high implementation fidelity on this matrix a program would need to receive at least 85% of the total available points or in this case at least 53 points. A moderate degree of implementation fidelity would be indicated by an overall composite score between 60% and 84% of the total available points (38 to 52 points out of a possible 63 on this matrix) and low implementation fidelity would be reflected in an overall composite score of less than 60% of the total available points (37 points or fewer out of a possible 63 on this matrix).

Background

Overview of participating districts

MA ESE recruited 18 districts to participate in the first year of the implementation of the Pyramid Model as part of the SSIP, May 2015 to September 2016. While these 18 districts are only a fraction of the 408 districts in Massachusetts, they educate more than 21% of the total number of public school students with and without IEPs enrolled in preschool through grade 12 in the Commonwealth. In addition, the 18 participating districts are responsible for educating 3,041 preschool children with IEPs or 28.3% of the total statewide preschool population with IEPs (Geldart, 2016). The students served by these districts are primarily educated in a fully inclusive environment (42% of students) or a partially inclusive environment (26%)(Geldart, 2016). Approximately 22% of students in the participating districts are educated in a substantially separate setting and the remaining 10% receive their special education services in the home or at a service provider's office. Similar to the statewide demographics, preschool students with disabilities in the 18 participating districts fall into three primary disability categories: developmental delay (46%), communication

(33%) and autism (21%), with the remaining 6% representing other low incidence disability categories. Figure 13 shows a map of all 18 districts participating in this initiative. They are geographically dispersed throughout the state of Massachusetts but a majority of them are located in the eastern half of the state. While many of the largest districts in the state are included in this sample, there are also several smaller districts serving around 2500 students in total and both rural and urban populations are represented. Lastly, the 18 districts represent communities of diverse economic means including ones that serve predominately families who qualify as low income and those that serve more affluent communities.

Qualifications and training for external coaches

In the first year of implementation MA ESE contracted with four individuals and two organizations to support the implementation of the Pyramid Model in the state. One organization employed four individuals who served as external coaches for this initiative while the other organization provided two individuals to support a train-the-trainer model in districts for Positive Solutions, the parent training modules included in the Pyramid. In total MA ESE funded eight external coaches, with each assigned to more than one of the 18 participating districts, to support their first year of implementation. As part of the procurement process each external coach submitted an application to MA ESE including a resume and was required to provide a sample one hour training to MA ESE staff. The prevalence of the Pyramid Model in Massachusetts supported MA ESE's ability to identify coaches with significant experience coaching in the Pyramid Model. However, because the adoption of the Pyramid Model has been much more prevalent in Head Start and child care programs, the external coaches had less experience working with public preschool programs. In addition, the external coaches had far more experience providing classroom coaching to

teachers implementing the model and less experience supporting district and school-wide leadership teams looking to implement at scale across a program.

External coaches were hired to serve as both coaches and trainers in the Pyramid Model practices. External Coaches were given the opportunity in the summer of 2015 to co-train with a national trainer provided by the Pyramid Model Consortium to demonstrate their ability to train large audiences in the preschool Pyramid Model practices. Evaluations from participants and ESE observations of these initial co-training opportunities indicated a need among most of the external coaches for additional training in how to tailor the training content to the unique needs public preschool teachers and how to delve deeper into the content with educators who have advanced degrees and/or previous exposure to the Pyramid Model practices. In addition, many of the external coaches had substantial experience providing training to small groups of educators but limited experience facilitating trainings with 100 or more participants.

From October 2015 to April 2016 MA ESE provided group and individualized training to the external coaches in collaboration with the national trainer. These training sessions included eight, one hour conference calls in which trainers participated in a group discussion about the content of the Pyramid Model practices training modules and how each section of the training could be tailored to the unique needs of Massachusetts educators. External coaches also co-facilitated at least one half-day of the four day Pyramid Model practices training and received feedback about their training from participants, the national trainer, and the other external coaches. External coaches who needed additional support also participated in coaching phone calls with the national trainer and MA ESE staff about their specific strengths and skills that needed further development. Throughout the study period

external coaches participated in calls and face-to-face meetings with the author to support their coaching activities with participating districts and trouble shoot challenges.

Overview of interview participants

MA ESE has contracted with the Pyramid Model Consortium to provide support to the state in implementing this model with fidelity and statewide training activities. The author, as part of the data collection activities for this study will conduct two interviews with individuals from the Pyramid Model Consortium who have been heavily involved in the work in Massachusetts. The first interviewee was Rob Corso, the current Executive Director of the Pyramid Model Consortium. Mr. Corso served as the Project Coordinator for CSEFEL, the organization which developed the Pyramid Model, from 2001 to 2013. Mr. Corso is also the primary contact on the current contract between MA ESE and the Pyramid Model Consortium to support implementation of the Pyramid Model in Massachusetts. Mr. Corso is currently working with 29 states to support their statewide implementation of this model.

The second interview was conducted with Kristin Tenney-Blackwell, a contractor with the Pyramid Model Consortium who provides Pyramid Model trainings and coaching on their behalf and has more than 16 years of experience in the field. Mrs. Tenney-Blackwell was very involved in the first year of implementation of the Pyramid Model in Massachusetts. She was the primary or co-presenter at each of the Pyramid Model Practices trainings offered by MA ESE in the identified time frame for this study. In addition, she provided significant coaching support to the MA ESE external coaches including group and individualized coaching calls to support the external coaches' ability to offer the practices trainings independently

Conclusion

The fidelity of implementation of the Pyramid Model in 18 districts in Massachusetts can be measured at three distinct levels: the state, the district, and the classroom. A process evaluation of this implementation examines the extent to which each of the activities outlined in the theory of action and articulated in SSIP are delivered as planned and consistent with the guidance from the developers of the national Pyramid Model in order to contribute to an overall analysis of implementation fidelity of the initiative. At each level of implementation there are key indicators of fidelity that can be measured and evaluated to inform a cycle of continuous improvement and ongoing refinement of the implementation plan to meet the needs of educators in Massachusetts and ensure the scalability and sustainability of this model at all levels of implementation.

Chapter 4

Results

Introduction to the Results Section

Following a brief overview of the participating districts, the results section is divided into two sections, the results of the study by data collection tool and the results by each of the five dimensions of fidelity: adherence, duration or dose, quality of delivery, participant responsiveness, and program differentiation. The section arranged by data collection tool provides an overview of each data collection tool and the study's results based on that tool. The second section introduces the results matrix found in Table 7 and the outcomes of this study based on the information gathered using the data collection tools and the definitions of high, medium, and low fidelity in each dimension found in the matrix.

Overview of Results by Data Collection Tool

The State Systemic Improvement Plan (SSIP). The implementation of the Pyramid Model in Massachusetts was part of broader statewide policy and practice changes associated with the SSIP. The Phase I SSIP Report was due to OSEP on or before April 1, 2015. The Phase I SSIP report had several required components including a comprehensive data analysis of extant state data, an analysis of existing infrastructure to support improvement and build capacity for improving child level outcomes, the identification of a state-identified measurable result (SIMR), the selection of a coherent improvement strategy or strategies, and a detailed theory of action (OSEP, 2014). The theory of change for the Massachusetts SSIP is presented in Figure 1. Each required component of the SSIP report contained multiple elements that were independently analyzed by members of the OSEP staff. OSEP staff members examined the quality of each component and the overall quality of the Phase I

report using an OSEP-created published evaluation tool to assess how well states met each of the required components based on the description provided by the state and the level of technical assistance a state would need using OSEP's Framework for Differentiated Monitoring and Support (OSEP, 2014).

MA ESE received feedback on the initial submission of the Phase I report in July of 2015 through a phone call with the Massachusetts OSEP liaison, the State Director for Special Education, and MA ESE Staff. This phone call indicated that the state Phase I report was approved without concerns. The OSEP staff praised Massachusetts for the detailed and comprehensive information provided in the Phase I report and the alignment between the Part B and Part C SSIPs in the state. In addition, the 2015 state determination letter from OSEP indicated that Massachusetts met the requirements and purposes of Part B of the IDEA based on the totality of the State's data and information provided. Information analyzed included the Federal fiscal year (FFY) 2013 State Performance Plan/Annual Performance Report (SPP/APR), additional State-reported data, and other information that is publicly available such as postings on the MA ESE website (Musgrove, 2015).

The Phase II SSIP report was due to OSEP on April 1, 2016 as part of the FFY 2014 SPP/APR. In addition to any updates of the Phase I report, OSEP required that the Phase II report include information on infrastructure development, planned support for district implementation of evidenced based practices, and an evaluation plan. MA ESE again received feedback on this submission through a phone call with OSEP staff and the resulting state determination letter. The 2016 state determination letter from OSEP indicated that Massachusetts again met the requirements and purposes of Part B of the IDEA based on the totality of the State's data and information. Information analyzed included the Federal fiscal

year (FFY) 2014 SPP/APR, additional State-reported data, and other information that is publicly available (Ryder, 2016). A phone call with the Massachusetts OSEP liaison, MA ESE staff, and the Massachusetts state special education director indicated that the Phase II SSIP report was accepted by OSEP without revisions.

Interviews. Two interviews were conducted with individuals from the Pyramid Model Consortium in the fall of 2016. Additional information about each individual is provided in the procedure section of this document. Mrs. Tenney-Blackwell provided input on the adherence, duration and dose, and quality of statewide training opportunities provided by MA ESE. Her feedback indicated overall high levels of alignment with the Pyramid Model as designed. In addition while MA ESE made some modifications to the training materials which are discussed in further detail later, she indicated that the content of the trainings was consistent with how the model was designed and that the structure and delivery of the content was also aligned with how the model was intended to be delivered. Lastly, Mrs. Tenney-Blackwell provided feedback on the training provided to external coaches which is addressed in more detail in the discussion section on external coaches.

With significant experience in implementing the Pyramid Model across multiple states and a longstanding relationship with Massachusetts, Mr. Corso provided information on the implementation of this model at the state level and a systems perspective. He indicated that the trainings were of high quality and used the recommended materials from the Pyramid Model consortium appropriately. He also noted that there were necessary modifications to the trainings to reflect Massachusetts' unique needs and the capacity of the trainers to deliver the content but that those did not impact the fidelity of the content. Lastly, he discussed concerns with the ability of some of the districts to move beyond classroom implementation

to program wide implementation of the Pyramid Model with the support of the external coaches. While MA ESE has put an appropriate emphasis on systems building and approaching this initiative from a perspective that develops locally sustainable practices, the skills and capacities of the individuals at the district level can impact districts' abilities to scale up implementation and adhere to the model. These challenges are detailed in the discussion section.

Training data. Training data collected as part of this study provides information on the scope and sequence of training activities during the study period. By examining the order of training activities, participant engagement as measured by the number of districts attending, and the revisions to the training schedule based on participant and national trainer feedback, this data provides information on the way in which MA ESE approached statewide implementation. This training data also informs an analysis of whether or not the execution of these activities aligns with the implementation plan detailed in the SSIP and the ways in which the developers of the model intended it to be presented according to the interviews described previously. Table 10 presents summary statistics of the trainings that occurred during the identified timeline including the name of the training, the length of the training, the number of times the training was offered, and the number of individuals and districts that participated in each type of training. In addition to the trainings outlined in the SSIP Phase I implementation plan, MA ESE identified the need for additional trainings for the external coaches. From May 2015 to September 2016 ESE provided individual and group coaching opportunities for the external coaches to develop their ability to provide statewide trainings in the Pyramid Model practices. These trainings were co-led by the author and a Pyramid Model Consortium trainer with expertise in the Pyramid Model. In addition, each external

coach had at least one opportunity to participate in the Pyramid Model Practices Training through a co-training model with the national trainer to further develop their presentation skills. Each coach participated in a private debriefing meeting or phone call with the national trainer and the author after they were co-trained.

An analysis of attendance at statewide training activities as presented in Table 10 indicates that a majority of districts participated in most trainings with no fewer than 14 of the 18 participating districts (78%) represented at any one training. Every district participated in the Initial Leadership Team Meeting and at least one Practices Training. Districts were least likely to participate in the End of Year Leadership Team Meeting that occurred in June of 2016 and the Beginning of the Year Leadership Team Meeting that occurred in September 2016. Participant feedback indicated that the scheduling of these meetings was particularly challenging given how close they occurred to the end and beginning of the school year. District leadership teams that were unable to participate in these trainings were asked to schedule a meeting with their external coach to review the content of the training at a later date. Districts that did not participate in the TPOT training in year one of implementation (three districts) either failed to complete the Pyramid Model Practices Training prerequisite training prior to the TPOT training or indicated a lack of staff availability for coaching in the first year of implementation and therefore delayed registration for the TPOT until year two. Overall, the scope and sequence of the trainings presented in Table 10 is consistent with the training delivery model designed by the creators of Pyramid Model according to the interviews with them described previously. In addition, the high participation rates among districts in this initiative indicates that the districts were generally accessing the training necessary to successfully implement this model.

District Benchmarks of Quality. The Pyramid Model Benchmarks of Quality is a self-assessment that identifies 47 benchmarks of quality in nine critical element categories on which programs can assess their implementation on a scale of not in place, partially in place, or in place. A copy of the Benchmarks of Quality can be found in Appendix B. Created by the Pyramid Model Consortium, this tool allows programs and districts implementing the Pyramid Model to assess their own progress over time and provides the state with data on programs' implementation status, including current challenges (L. Fox, personal communication, March 15, 2015). Each of the 18 participating districts completed the Benchmarks of Quality at one of the three initial statewide Leadership Team Meetings which occurred in May 2015, September 2015, and October 2015. Districts reassessed their implementation of the model at the Midyear Leadership Team Meeting in March 2016 and again at the End of Year Leadership Team Meeting in June of 2016. The aggregated results across all participating districts for the beginning, mid-, and end-of-year administrations of the Benchmarks of Quality for the first year of implementation in Massachusetts can be found in Figure 13. However, not all 18 districts participated in each of these meetings and therefore results from a beginning, middle, and end of year administration of the Benchmarks of Quality are not available for each district. Figure 13 includes all available beginning, middle, and end-of year Benchmarks of Quality and therefore includes 18 districts in the beginning of the year data, 11 districts in the middle of the year data, and 14 districts in the end-of-year data. Overall 12 of the 18 districts completed beginning, middle, and end-of year Benchmarks of Quality. Two districts participated in the beginning of the year administration only. Two districts completed the beginning and midyear Benchmarks of Quality but not the

end-of-year administration and two districts completed the beginning and end-of-year Benchmarks of Quality but not the mid-year administration.

Results from the Benchmarks of Quality completed by the 18 participating districts at the Beginning, Mid and End of Year Leadership Team Meetings are shown in aggregate in Figure 13. What follows is a description in the changes in the benchmarks from beginning, to the mid-year, to the end of year administration. At the beginning of the year administration participating districts indicated that 52% of benchmarks for the "Establish Leadership Team" critical element were not in place. In contrast, at the end of year administration participating districts indicated only 5% of benchmarks were not in place, 32% were partially in place and 63% of the benchmarks were in place. This represents the greatest shift from the beginning of the year to the end of year across the nine critical elements.

Participating districts indicated that only 15% of benchmarks for "Staff Buy-in" were in place at the start of this initiative. However, by the end of year administration 61% were in place. While this is a substantial shift over the first year of implementation it does indicate there might be some challenges with participant responsiveness to implementation. The third critical element on the benchmarks of quality is "Family Involvement". The participating districts showed the least progress on this critical element with 5% of benchmarks in place at the beginning of the year and 9% in place at the end of the year. Districts in this initiative made significant progress in implementing "Program-wide Expectations," the fourth critical element. While 52% of benchmarks were not in place at the beginning of the year, 60% were in place by the end of year administration. There was less of a shift in the results for the "Strategies for Teaching and Acknowledging the Program-Wide Expectations" critical element from the beginning to the end of the year. At the beginning of the year 20% of

benchmarks were in place while at the end of year only 29% of benchmarks were in place for this critical element.

The critical element "All classrooms demonstrate the adoption of the 'Teaching Pyramid'" had a large percentage of benchmarks (74%) partially in place at the beginning of the year benchmarks administration. By the end of year administration 49% of benchmarks were partially in place and 43% were in place, representing the largest percentage across the nine critical elements that were either partially or fully in place. Districts were also likely to report significant progress on the "Procedures for Responding to Challenging Behavior" critical element. Only 35% of benchmarks were fully in place at the beginning of the year leadership team meeting while 58% were in place by the end of the year. Districts were less likely to report progress on the "Staff Support Plan" critical element, only 26% of benchmarks were in place by the end of year administration. Lastly, districts reported some of the lowest progress on the "Monitoring Implementation and Outcomes" critical element. There was only a eight percentage point increase in the percentage of benchmarks partially in place over the course of the year (23% to 31%) and only a ten percentage point increase in the percentage of benchmarks fully in place (0% to 10%).

Overall the Benchmarks of Quality completed at the Midyear Leadership Team Meeting reflected some self-correction by leadership teams as they became more familiar with the model. External coaches reported that leadership teams rated more items as "not in place" or "partially in place" after participating in additional Pyramid Model trainings and therefore the aggregate results on some of the critical elements reflects lower implementation than was indicated on the initial Benchmarks of Quality. For example, the "family involvement" benchmarks went from 5% of benchmarks in place at the beginning

administration to 2% of benchmarks in place at the midyear administration. While these aggregate results represent meaningful change for the districts that completed the benchmarks of quality at least two different times, the midyear administration results are missing Benchmarks of Quality from seven of the participating districts and the end of year results are missing Benchmarks of Quality from four participating districts. Therefore, these results may not represent the totality of the changes or lack thereof across the entire group of participating districts.

TPOT. The Teaching Pyramid Observation Tool for Preschool Classrooms (TPOT) is a published coaching and classroom implementation fidelity tool for the Pyramid Model (Hemmeter, Fox, and Snyder, 2013). According to the authors, the TPOT utilizes a two hour classroom observation and subsequent interview with the teacher to score implementation of key Pyramid Model practices in 14 areas, identify the presence of *red flags* during instruction or in the classroom, and understand an educator's response to challenging behavior. The original implementation plan for the Pyramid Model in Massachusetts asked that implementing districts complete a TPOT for each teacher implementing the model at least twice a year (MA ESE, 2015). The implementation plan outlined the intention that the TPOTs would occur shortly after teachers were trained in the model and that aggregate data on the use of the TPOT with implementing teachers would be submitted at least annually to MA ESE. However, a number of challenges occurred when trying to implement this plan.

Statewide, in the initial year of implementation, 59 individuals in Massachusetts were trained and tested as reliable in the TPOT as part of this initiative. According to the implementation plan, districts were to identify at least one or two individuals who could serve as classroom-based coaches in the Pyramid Model who could also administer the

TPOT as part of their coaching activities. In practice, districts expressed significant challenges with identifying these coaches primarily due to staffing shortages and conflicts with existing union contracts. According to registration data and reports from the external coaches, a majority of TPOT training participants were individuals in administrative or supervisory roles who were interested in learning more about the tool but who did not have the time or ability to administer the TPOT in full with their teachers. In addition, while many individuals had been trained in the Pyramid Model Practices, many of the participating programs did not plan on officially launching the model in their districts until the 2016-2017 school year. Therefore, while these teachers had been trained, district personnel did not consider them to be implementing the model prior to the official district-wide launch of the initiative and did not conduct TPOTs with these teachers in the 2015-2016 school year.

As a result of these challenges, MA ESE decided not to collect aggregate TPOT scores in the initial year of this project, the 2015-2016 school year. Therefore these data are unavailable for analysis. Beginning in the 2016-2017 school year MA ESE will collect data on the number of TPOTs conducted but not the results of those TPOTs. External coaches will support the classroom-based coaches in the collection and analysis of this data at the district level. In addition, the Phase III SSIP report, which will be submitted to OSEP on April 1, 2017, will reflect this change in the evaluation plan.

External coach contact records. The external coach contact record was designed to be completed following any substantive contact between a MA ESE-funded external coach and a participating district including planning phone calls, leadership team meetings, and other site visits. The first year of implementation required district personnel to attend off-site trainings for up to twelve days over the course of the school year. This meant that it was

challenging for external coaches to find time to also conduct in-district site visits. From September 2015, when the external coach contact record form was created, to September 2016 only 21 records were completed by external coaches. Of those 21 records, almost all of them were missing responses to at least one or more of the questions. In addition, of those 21 records, 12 were for non-substantive contacts with the participating districts such as sending an email reminder or leaving a voicemail. Seven of the records detailed participation in statewide trainings that did not require the form to be completed. The two records addressing substantive contacts were incomplete. Therefore, for the timeline covered by this study, the external coach contact records were not a reliable source of data on the implementation status of participating districts. Data collection activities for the external coach contact records were supplemented by ongoing communication with the coaches as MA ESE-funded vendors. The author had significant contact with these coaches through face-to-face meetings and the statewide and external coach training sessions described previously and was therefore able to collect qualitative data on the current implementation status of participating districts through these conversations.

Results by Dimension of Fidelity

Table 7 presents a matrix of the five dimensions of fidelity as established by Dusenbury et al., (2003) and O'Donnell (2008) and each of the tools identified previously for assessing implementation fidelity of the Pyramid Model. For each tool and relevant dimension of fidelity the author has established criteria for high, moderate, and low fidelity and the points assigned to each of these ratings (3, 2, and 1 respectively). The score awarded to this initiative in the table is indicated by the response in italics within the table. Table 8 presents the number and percentage of times each tool was rated as having high, moderate,

and low implementation fidelity across the dimensions of implementation fidelity. Only the training data was ranked as having high implementation fidelity 100% of the time. The SSIP Reports were rated as having high implementation fidelity 75% of the time while the interviews were identified as having high implementation fidelity 67% of the time. The TPOT had high implementation fidelity on only one factor (25%) and was the only tool to be rated as having low implementation fidelity with 3 factors (75%) rated as low. The district Benchmarks of Quality were rated as moderate implementation fidelity 100% of the time. The next several sections will review the results of this study on each dimension of fidelity which are summarized in Table 9 which shows the number and percentage of tools at high, moderate, and low fidelity for each relevant dimension of fidelity.

Adherence. Adherence is the extent to which the essential components of an intervention are implemented as intended (Dusenbury et al., 2003; O'Donnell, 2008). An examination of information from the SSIP reports, interviews, training data, district Benchmarks of Quality, TPOTs, and external coach contact records indicates an overall score of 13 out of a possible 18 points (72% in the moderate range) on the adherence dimension of fidelity. On the adherence dimension MA ESE achieved high implementation fidelity for two tools (the SSIP reports and training data) or 33.3% of tools related to this dimension of fidelity; moderate implementation fidelity for three tools (interviews, district benchmarks of quality, and external coach contact records) or 50% of tools related to this dimension of fidelity, and low implementation fidelity for one tool (the TPOT) or 16.7% of tools related to this dimension of fidelity.

MA ESE achieved high fidelity of adherence according to the information analyzed in the SSIP reports and the training data. MA ESE achieved moderate adherence fidelity

according to the interviews which indicated that there were minor variations in the delivery of the content by external coaches. MA ESE was also awarded "moderate adherence" based on the district Benchmarks of Quality and the external coach contact records. The most recent Benchmarks of Quality for all participating districts which was completed in June 2016 indicated that 40% of the total number of benchmarks are fully in place (range of 9% to 63% across the nine critical elements) and 40% of benchmarks are partly in place (range of 31% to 60% across the nine critical elements). The external coaches reported some concerns with adherence to the model in four participating districts who indicated that they have only begun partial implementation of the practices in their classroom. This is reflected in the Benchmarks of Quality as well as approximately 12% of the practices benchmarks which were reported as not in place and 60% were only partially in place at the End of Year Leadership Team Meeting. As discussed previously, participating districts did not collect and submit TPOT data for implementing teachers in the identified timeframe and therefore MA ESE received a low level of adherence on this tool. Overall, these results indicate a moderate level of adherence fidelity (72%) which will be addressed in further detail in the discussion section.

Duration or dose. Dusenbury et al., (2003) and O'Donnell (2008) define the measurement of dose and duration of an intervention as the amount of program content received by participants. An examination of information from the SSIP reports, interviews, training data, district Benchmarks of Quality, TPOTs, and external coach contact records indicates an overall score of 11 out of a possible 18 points on the duration or dose dimension of fidelity or a moderate implementation fidelity score of 61%. On the dose or duration dimension MA ESE achieved high implementation fidelity for three tools (60%) and low

implementation fidelity for two tools (40%). MA ESE achieved high duration or dose fidelity according to the information analyzed in the SSIP reports, interviews, and training data. MA ESE achieved low duration or dose fidelity on the TPOT and external coach contact records. In the identified time frame the TPOT was completed by internal coaches for teachers implementing the model in less than 50% of participating districts. In addition, the external coaches reported providing on-site and/or virtual coaching at least every other month for less than 50% of the participating districts. Overall, these results indicate a moderate level of duration or dose fidelity (61%) which will be addressed in further detail in the discussion section.

Quality of delivery. Quality of delivery is the extent to which "the implementer delivers the program using the techniques, processes, or methods prescribed" (O'Donnell, 2008, p. 34). An examination of information from the SSIP reports, interviews, training data, TPOTs, and external coach contact records indicates an overall score of 14 out of 15 points on the quality of delivery dimension of fidelity or 93%, a high level of implementation fidelity for this dimension. On the quality dimension MA ESE achieved high implementation fidelity for four tools (80%) and moderate implementation fidelity for one tool (20%). MA ESE achieved high quality of delivery fidelity based on information from the interviews, training data, TPOT, and external coach contact records. The SSIP reports indicated moderate quality of delivery fidelity as it reflected minor modifications in the methods and resources used to deliver the intervention.

This was the only dimension of implementation fidelity for which MA ESE achieved an overall high level of fidelity with a score of 93% which will be addressed in further detail in the discussion section.

Participant responsiveness. Participant engagement in an intervention is a critical element for success (O'Donnell, 2008). Participant responsiveness refers to the degree to which individuals are involved in and utilizing the components and activities of an intervention (Dusenbury et al., 2003; O'Donnell, 2008). An examination of information from the SSIP reports, training data, district Benchmarks of Quality, TPOTs, and external coach contact records indicates an overall score of 11 out of 15 points (73%) or a moderate implementation fidelity score on the participant responsiveness dimension of fidelity. On the participant responsiveness dimension MA ESE achieved high implementation fidelity for two tools (40%), moderate implementation fidelity for two tools (40%), and low implementation fidelity for one tool (20%). MA ESE achieved high participant responsiveness fidelity based on information in the SSIP reports and the training data. MA ESE achieved moderate participant responsiveness fidelity based on the Benchmarks of Quality which were completed by more than 50% but less than 80% of participating districts and the external coach contact records which indicated that more than 50% but less than 80% of districts were actively engaged in implementation activities in classrooms and across the program/district during the identified timeframe.

Overall, these results indicate a moderate level of adherence fidelity (72%) which will be addressed in further detail in the discussion section.

Program differentiation. An examination of the fidelity of implementation through the lens of program differentiation is intended to establish whether or not the critical features of a program that set it apart from a control condition are present or absent as part of an implementation. The implementation of the Pyramid Model by MA ESE did not include a

control condition and therefore there is no information on the level of program differentiation fidelity for this study.

Summary of results. The matrix established bands of summary scores across five dimensions of fidelity for high, moderate, and low implementation fidelity. In order to achieve a composite score of high fidelity for this study MA ESE would need to receive a score of at least 56 points or 80% of the total available points (66 points). Based on the analyses presented here MA ESE achieved an overall score of 49 points (78%) representing a moderate level of implementation fidelity (within the range of 39-55 points). MA ESE had the highest degree of fidelity on the quality of delivery dimension (14 out of 15 possible points) and the lowest degree of fidelity on the duration or dose dimension (11 out of 18 possible points). Factors contributing to these results are discussed in the next section.

Chapter 5

Discussion

Summary of Findings and Conclusions

The implementation of the Pyramid Model in Massachusetts as part of the SSIP achieved an overall rating of moderate implementation fidelity using the matrix. While MA ESE did not achieve an overall composite rating of high implementation fidelity, as shown in Table 8, each of the data collection tools were assigned at least one rating of high implementation fidelity with the exception of the district Benchmarks of Quality which received two ratings of moderate implementation fidelity. Table 9 presents the number and percentage of tools at high, moderate, and low implementation fidelity. For each of the four assessed dimensions of fidelity MA ESE achieved a rating of high implementation fidelity on at least two tools. As shown in Tables 5 and 9, MA ESE was the least likely to have tools rated as high for implementation fidelity of the adherence dimension with 33.3% of relevant tools rated as high and the most likely to have tools rated as high implementation fidelity for the quality dimension which had 80% of relevant tools rated as high. MA ESE was most likely to have a tool rated as low implementation fidelity on the dose or duration dimension (40% of relevant tools were rated as low implementation fidelity) and the participant responsiveness dimension (20% of relevant tools were rated as low implementation fidelity) The next section begins with a review of implications for professional practice as these are the factors that most likely contributed to the overall rating of moderate fidelity for this initiative. The discussion that follows describes the practical considerations for other researchers or states seeking to implement a similar statewide initiative and the limitations of this study.

Implications for professional practice.

Identification and recruitment of participating districts. In addition to the selection criteria described previously that were used to identify participants for this initiative, MA ESE may have contributed to a lack of readiness for implementation among participating districts as a result of its participant identification and recruitment procedures.

Districts that were offered an opportunity to join the first training cohorts in 2015 were contacted individually for a discussion about the requirements for participation and the resources available from MA ESE to support their implementation. As part of this process districts were asked to complete a "Readiness Checklist" that articulated the requirements for successful implementation of the Pyramid Model and required signatures from multiple administrative and educational leaders in the district, including the special education director and the superintendent, to demonstrate buy-in from all levels of administration. However, because the initiative was still in development, the "Readiness Checklist" provided limited information about the scope of the initiative. It did not provide extensive details about the number of days of training required to participate in the initiative, expectations about data collection requirements, and a clear description of what program-wide implementation of the Pyramid Model looks like in a district. Therefore, while MA ESE ensured buy-in from administrators, it did not clearly articulate exactly what they were agreeing to participate in as it had not been clearly defined by MA ESE itself.

Guidance from OSEP regarding the SSIP recommended that implementation of the evidence-based strategy for the SSIP begin late in Phase II (2015-2016 school year) or at the beginning of Phase III (2016-2017 school year) of this work (OSEP, 2013). However, there was strong administrative support in MA ESE to begin implementation as soon as possible and therefore

recruitment for participants began before the Phase I report was finalized and implementation began less than 2 months after the Phase I report was submitted to OSEP and prior to receiving feedback from OSEP on the plan included in the Phase I report. This rush to implementation meant that districts were committing to participate in a multi-year initiative without a finalized plan or a full schedule for year one activities, including identifying which trainings would be required. As a result, as implementation progressed from May 2015 to September 2016, the expectations and requirements of this initiative were refined and in some cases expanded which presented challenges for the districts. Districts reported that they were required to spend more time out of district than initially anticipated, were required to collect more and different data, and faced problems identifying and training an internal, classroom-based coach due to union restrictions and staff availability. In addition, while districts participating in this initiative were offered an additional $2,000 on a state allocation grant, the timing of the initial trainings meant that most funds outside of this small additional allocation had already been encumbered and therefore were unavailable to support the year one activities.

The size of participating districts. As described in the results section, while the participating 18 districts represent a small percentage of the total number of districts in the state, the participant list includes seven of the 15 largest districts in the state. The size of these participating districts meant that they were more easily able to identify personnel and resources to support their participation. Their larger size may also have allowed them to identify higher performing teachers to participate in the initial stages of the initiative. However, it also meant that there was significantly more work involved in creating a structure to support district-wide implementation of the Pyramid Model. District leadership

teams from larger districts in some cases consisted of individuals who worked at different sites and lacked common planning time, a challenge for implementation. In addition the larger districts had significantly more staff to train in order to achieve program wide implementation. While the large districts frequently had individuals designated as classroom coaches for other initiatives, these coaches often already had a full coaching load and reported it was a challenge to incorporate Pyramid Model coaching activities into their work. In contrast, smaller participating districts were more likely to have leadership teams who were already working closely together. However, these smaller districts faced challenges identifying substitutes from their limited pool to enable classroom teachers to attend training.

The structure of the preschool program. When recruiting districts, MA ESE did not take into consideration the structure of the preschool program in a district. The 18 participating districts' preschool programs represent two different program models. Approximately two thirds of the districts have preschool classrooms spread across a number of elementary school buildings. In these cases the preschool class or classes are part of a larger building serving students up to fifth or eighth grade. In contrast, about one third of participating districts utilize an early learning center model in which all of the preschool classrooms are housed in the same building and that building primarily or only serves preschool students. An examination of the implementation data disaggregated by preschool program structure indicates that single site districts were able to being implementation earlier and more easily implement systems to support teachers' use of the Pyramid Model including establishing program-wide expectations. Multi-site districts may have had more challenges maximizing the contextual fit between the Pyramid Model implementation and each unique site, a critical feature for scaling up evidence-based practices. Given the mix of preschool

program structures in the state it is appropriate to have both models represented in the initial implementation cohort. However, the significant differences in challenges faced by programs depending on their structure should be shared with new participants. As MA ESE seeks to expand implementation into additional districts it may also be beneficial to pair new districts with a similarly structured district from the initial cohort of 18 so that they can share challenges and strategies for success.

External coaches. As described in the Procedure section, the External Coaches required a great deal of additional training and support that was not anticipated by MA ESE at the time they were contracted to provide coaching services. These additional professional development needs required a significant portion of the budget allocated for the contract with the Pyramid Model Consortium and the contracts with the external coaches as they were paid to participate in these training activities. These additional training needs also meant that the external coaches were not able to participate in many face-to-face coaching activities with their assigned districts. In the first year of implementation external coaches should have participated in at least four site visits with their district leadership teams in addition to the statewide face-to-face trainings and leadership team meetings. However, due to these additional training commitments and a limited budget for the year, none of the 18 participating districts had more than two site visits and a majority only had one from their external coach.

The structure of the first year of training. As described previously, MA ESE launched this initiative before finalizing the structure of the first year of training and coaching activities. The scope and sequence of activities as it was delivered is presented in Table 6. As part of the first year of this initiative district leadership team members were

asked to attend a Beginning, Mid, and End of Year Leadership Team Meeting. In addition, they were strongly encouraged to attend a Pyramid Model practices training to learn about the practices teachers would be expected to implement. Districts were asked to include a classroom-based coach and individual with behavioral expertise on their leadership team. Each of these individuals were encouraged to attend a Prevent, Teach, Reinforce Young Children (PTR-YC) training and a TPOT training, both of which had the Pyramid Model Practices training as a prerequisite. Therefore, some leadership team members were potentially being asked to participate in more than 11 days of off-site training in a single school year. The final schedule for the entire year of trainings was not completed until after the Beginning of the Year Leadership Team Meeting in September 2015. In addition, due to scheduling issues, some training dates had to be shifted mid year which then created conflicts with other in-district activities and trainings. Lastly, the significant number of trainings out of district required districts to identify substitutes and pay for travel costs for participants which further taxed a limited district budget for this initiative.

The physical location of the statewide trainings also presented a challenge. Due to procurement restrictions, MA ESE only had three possible locations from which to choose a training venue. All statewide trainings were held at the most central location available which was in Marlborough, MA, a town about one hour west of Boston. However, because the 18 participating districts represented all regions of the state, many teams were required to travel two hours or more each way to attend trainings in Marlborough. Some teams had to travel more than three hours each way and were required to spend some of their limited funding to pay for overnight stays in Marlborough for multi-day trainings. The geographic location of

the trainings limited the ability of some districts to send their full leadership team to the trainings and their ability to send teachers to the practices trainings.

Pyramid Model practices trainings. The Pyramid Model practices trainings consist of four modules developed by CSEFEL: module 1 which covers the bottom of the Pyramid including building relationships and creating supporting environments; module 2 which covers explicit social emotional teaching strategies, and modules 3A and 3B which address intensive individualized interventions including developing behavior support plans (CSEFEL, 2015). Initially MA ESE followed the recommendation of the Pyramid Model Consortium and offered these trainings as a four day series with one day allocated for each module. However, feedback from the first two sets of four day practices trainings indicated that Massachusetts educators felt the trainings were too long and had too much of a review of teacher's existing knowledge without introducing new content and strategies. Upon consultation with the experts from the Pyramid Model Consortium it was determined that the longstanding presence of the Pyramid Model in Massachusetts had led to the penetration of Pyramid Model practices in many classrooms in the state without explicit training. In addition, the modules developed by CSEFEL were designed to be used with all educators, including those working in child care or other settings who may not hold a post-secondary degree. A majority of the participants in MA ESE's practices trainings however held masters or other advanced degrees with significant training in working with students with disabilities. Upon review of this information it was decided that the practices trainings needed to be tailored to the Massachusetts context including incorporating more in-depth content and advanced strategies that build upon those presented in the Modules. MA ESE worked closely with a consultant from the Pyramid Model Consortium to modify the training content and

shorten the training from four to three days. Interviews with the creators of the model and the national trainer confirm that the fidelity of the content was not altered but expanded and deepened to reflect the state's needs. However, while the national trainer received strong positive reviews for her presentation of this new content, several of the external coaches struggled with presenting the more nuanced and advanced content which lead to the need for additional training from the national trainer and MA ESE staff as described previously.

For the second year of implementation (2016-2017) MA ESE made significant changes to the structure of statewide training activities as a result of the challenges and feedback received in the first year of implementation. The most significant change was to the structure and location of the Pyramid Model practices trainings. MA ESE issued a request for quote to the current external coaches and procured one of the vendors to provide four regional practices trainings over the course of the 2016-2017 school year. This contract requires that the vendor provide two trainings in the eastern part of the state, one in the fall and one in the spring, and two trainings in the western part of the state, one in fall and one in the spring. In addition, the vendor was asked to identify host districts for each training and to schedule the training activities according to the needs of the host district. The goal of this structure is to minimize time out of district and travel time for participants and to support the unique needs of the participating districts. MA ESE also eliminated the Midyear Leadership Team Meeting to provide additional time for in-district support from external coaches. In addition, MA ESE added a coaches training day at the beginning of the current school year in September 2016. The morning of that daylong training was designed to build connections among classroom based coaches both within and across districts and to support improved data collection and family engagement activities based on the results of the end of year

Benchmarks of Quality. The second half of the day was devoted to an external coaches meeting in which challenges, opportunities, and expectations for the 2016-2017 school year were discussed with the coaches, MA ESE staff, and the national trainers. Lastly, MA ESE established the training schedule for the 2016-2017 school year early in the summer of 2016 so that districts had ample time to plan for statewide training activities.

Lack of on-site support and activities. A critical component of MA ESE's theory of action for this initiative was the use of external coaches to provide on-site support to district leadership teams and classroom-based coaches. However, a number of factors limited the external coaches' ability to provide this support in the first year of implementation. First, the significant number of days out of district for required Pyramid Model trainings meant that district leadership teams had few if any professional development days to offer for in-district support. Rather than holding regular, self-sustaining leadership team meetings in-district that were supported by occasional statewide activities, leadership teams were required to devote significant time to statewide activities leaving less time for more meaningful work in their program. In addition, MA ESE paid external coaches for attendance at statewide training activities which strained an already limited budget for their first year of work given the number of days of training they attended. Several external coaches were not hired until partway through the first year of implementation so some of the districts had their coach changed in the first six months of the project challenging existing relationships. In addition, the timing and location of statewide activities limited leadership teams' participation and therefore further limited the ability of the external coaches to develop strong working relationships with the teams that they were supporting. This time out of district also contributed to the inability of classroom-based coaches to conduct TPOTs for participating

teachers. This limited the ability of MA ESE to quantify the level of implementation of Pyramid Model practices in implementing classrooms and may have impacted district leadership teams' understanding of how the model was being implemented in classrooms within their district. External coaches began co-administering the TPOT with classroom-based coaches in the 2016-2017 school year to build additional capacity among classroom-based coaches for this tool and to support the use of the TPOT more broadly. As this initiative expands and TPOT data becomes available it will be valuable to analyze the relationship between a district's TPOT scores and their child outcomes data.

Most of the trainers procured by MA ESE were also located in the Boston metropolitan area while the participating districts are dispersed across the state. This further contributed to challenges scheduling face-to-face meetings with district leadership teams. Lastly, a great deal of emphasis in the first year was spent on building the external coaches' capacity to provide the practices training as described previously. This focus was at the expense of supporting their work in-district with leadership teams to further implementation of the Pyramid Model. Conversations with external coaches and the external coach contact records indicated that there was some confusion about their role and expectations regarding the frequency of face-to-face coaching. In the 2016-2017 school year MA ESE took additional steps to clarify these expectations including jointly developing an external coach expectations document with the coaches that outlined expectations for them and for MA ESE.

Generating buy-in at the local level. The limited availability of information on the expectations for participation in the initiative during the recruitment process and challenges in the first year of implementation lead to a lack of buy-in at the local level that may have

impacted district's ability to implement with fidelity. During the recruitment stage, MA ESE had not finalized the data collection requirements for participation, the scope and sequence of the trainings, or expectations for participation in years two and beyond. In addition, most of the participating districts were balancing participation in the Pyramid Model initiative and the statewide PBIS academies, each of which had their own requirements and expectations for implementing positive behavioral supports in a school.

Participating districts received a small grant of $2,000 to support their work on the Pyramid Model. However, in the future MA ESE can generate additional buy-in from potential districts during the recruitment process through several activities. First, MA ESE should ensure that the appropriate administrators are involved in the recruitment process and agree to the conditions for participation. Superintendents, special education directors, and building principals should be informed of the expectations for participation, the benefits of participation, and indicate their support through a commitment of staff time, fiscal resources, and political will.

Second, MA ESE should develop additional resources articulating the expectations for data collection activities for participating districts. In the first year of implementation districts were required to collect and submit several data points to MA ESE and the participating districts were very appreciative when this data was given back to them in a format that they could use to support their work. However, MA ESE did not clearly articulate the purposes and functions of some data collection activities which may have led to noncompliance with the request or only a partial response. Participating districts should be explicitly told what data will be collected, why it is important, how it will be used by the state, and most importantly how it can be used to improve their practice.

Third, MA ESE must continue supporting collaboration between the providers of the PBIS Academies and the Pyramid Model trainings. Participation in the PBIS academy is by school whereas participation in the Pyramid Model initiative is by district. Some schools participating in the PBIS academies are also a part of the Pyramid Model Initiative which has led many to create two separate leadership teams, one for schoolwide PBIS and one for the Pyramid Model. These two teams limit the ability to align support across the grade span and places an additional burden on staff time. This structure has also led many participating districts to prioritize the schoolwide PBIS initiative over the Pyramid Model initiative since the PBIS initiative impacts all grades rather than just preschool. MA ESE could generate additional buy-in for the alignment of the two programs by supporting additional collaboration between the national PBIS providers and the Pyramid Model Consortium. In addition, MA ESE could also develop new resources, in addition to the already disseminated crosswalk of the similarities and differences between the two models, that articulate how both the Pyramid Model and PBIS can be implemented in the same school and serve as complementary programs to support the social emotional development of students.

Lastly, MA ESE needs to structure the first year of participation in such a way that it minimizes the time district personnel need to spend out of district and maximizes the amount of time that external coaches can spend in-district supporting implementation planning. In the first year of this initiative MA ESE focused on statewide training activities at the expense of on-site individualized support. As a result districts who were already struggling with a myriad of issues reported that they felt overwhelmed with the requirements of the Pyramid Model and that they did not know where to begin with implementation. This sense of confusion and/or frustration may have contributed to a lack of buy-in at the local level and

delayed successful implementation until they could receive more individualized support from their external coach.

Utilizing additional implementation science resources. As described previously, the requirements of the SSIP established by OSEP were based on implementation science. The National Implementation Research Network (NIRN) has developed two resources that could support other states or large public agencies seeking to undertake a similar initiative. The Stages of Implementation Analysis Tool would have been particularly helpful as MA ESE proceeded with their implementation (NIRN, 2017). MA ESE began implementation of the Pyramid Model during Phase I of the SSIP (Spring 2015). However, according to the requirements of the SSIP as outlined by OSEP, states were to begin implementation of their evidenced based practice during Phase III which was scheduled to begin at the start of the 2016-2017 school year (OSEP, 2013). This rush to implementation led to several of the challenges described above. The Stages of Implementation Analysis Tool is a self-assessment that allows a program to analyze which aspects of each stage of implementation are in place, have been initiated or are partly in place, or are not in place and could have led to a more reflective, and ultimately slower implementation process for MA ESE. The authors of this tool identify the following stages of implementation: exploration, installation, initial implementation, and full implementation. MA ESE could use this tool to assess previous implementation activities, identify necessary changes to achieve full implementation, and to communicate with stakeholders on the successes and challenges facing implementation of the Pyramid Model. Other programs or states may be able to use this tool prior to implementation to analyze their existing infrastructure and to identify areas for improvement prior to moving to participant recruitment.

The second NIRN developed tool that may be useful to other organizations seeking to support a broad scale implementation of an evidenced based practice is the Hexagon Tool developed by Blase, Kiser, and Van Dyke (2013). The Hexagon Tool was designed to help states, districts, and schools systematically evaluate existing and new interventions. The tool uses six broad factors to inform this evaluation: needs, fit, resource availability, evidence, readiness for replication, and capacity to implement. This tool closely aligns with the infrastructure analysis activities required under the SSIP (OSEP, 2013). As MA ESE seeks to recruit additional districts this tool could be particularly helpful in defining and articulating readiness criteria for potential participating districts, including their current capacity for implementation. The tool includes criteria to identify whether or not appropriate staff are in place, there is sufficient buy-in at the district level, and whether or not resources are in place that can support sustainability. As described previously, the readiness checklist utilized by MA ESE during the recruitment process did not sufficiently articulate what would be required of participants. It also did not convey participation requirements beyond the first year of implementation. The Hexagon Tool can serve as a helpful guide in considering and outlining what is required for participation in an initiative and ensuring that those factors are in place prior to participation.

Alignment with community child care and Head Start. As described previously, many of the children attending public preschools in Massachusetts, including children with disabilities, also attend a community child care or Head Start program for at least a portion of their day. The literature review section of this study reviewed the importance of alignment of expectations and supports across environments for young children and especially for young

children with disabilities or challenging behaviors. Participants in MA ESE's Pyramid Model practices trainings indicated a familiarity with the Pyramid Model practices prior to their participation in trainings. However, the district leadership teams and external coach contact records reflected a concern among participating districts that a more significant need for support and training among community child care and Head Start program staff than public school staff. Recognizing this need MA ESE has worked closely with the Pyramid Model State Leadership Team which works with community-based programs to identify additional opportunities for training for these programs to support consistency across environments. In addition, MA ESE has encouraged district leadership teams to include representation from community child care and Head Start programs they work closely with in their regular leadership team meetings. MA ESE has allowed community child care and Head Start educators from programs working closely with participating districts to attend practices trainings and has opened up the TPOT trainings to community-based partners who have training in the pyramid model so that they can support coaching activities for high-fidelity implementation in their programs.

The Results Matrix as a Tool for Assessing Statewide Implementation Fidelity of an Evidenced Based Practice

Each state and territory is now required to develop a multi-year SSIP for their Part B and Part C special education programs and this trend is likely to continue as the federal government looks to states to demonstrate the results they have achieved for students, including those with disabilities, and not simply compliance with the law. This new SSIP requirement from OSEP for is part of a much larger push towards results driven accountability which is likely to continue (OSEP, 2014). The results matrix is potentially a

very valuable, customizable tool that could be used by states seeking to evaluate their implementation of an evidence based practice. Three critical components of the reporting requirements for the SSIP are the requirements that states implement an evidenced based practice, develop of a comprehensive evaluation plan, and engage with stakeholders on an ongoing basis. States implementing a SSIP or any comprehensive, multi-level statewide initiative, need a tool that can easily frame the scope of their data collection activities, communicate their expectations for what high fidelity of implementation looks like, and do so in a way that is easily understood and interpreted by stakeholders. Consolidating activities and data that occur at the state, district, classroom, and child level and evaluating them as a whole can be particularly challenging. While the results matrix presented here is customized to Massachusetts' implementation of the Pyramid Model, it can be repurposed to meet each of these needs and utilized assess the fidelity of implementation for any statewide initiative on the five dimensions of fidelity established by O'Donnell (2008).

The requirements for the SSIP issued by OSEP did not have any accompanying funding for program evaluation support. State education agencies need an easy and manageable way to identify the data that they need to collect and to identify the criteria by which those data collection activities will be evaluated. In order to utilize this tool to evaluate a statewide initiative a state team must first identify the key tools that will be used for data collection. In the case of the Pyramid Model, MA ESE used the SSIP reports (including feedback from OSEP), interviews with developers of the evidenced based practice, training data, the district Benchmarks of Quality, TPOT data, and the External Coach Contact Record tool developed by MA ESE. Most states implementing an evidenced based practice will likely have access to at least one tool created to assess implementation fidelity of the

evidenced based practice, experts in the practice being implemented, and training or coaching data related to their implementation in the state. Depending on the model used for implementation, states may also have access to other records of implementation that may be informative and can be incorporated into the matrix. Lastly, states implementing any mandated improvement strategy, be it an SSIP or other accountability measure, will likely have access to reports and feedback from stakeholders such as OSEP on their progress in meeting identified targets.

After identifying what they believe are the appropriate tools that will be used to assess implementation fidelity on all five identified dimensions, a state can then solicit stakeholder input on which tools inform which dimensions of fidelity and possibly work with stakeholders to identify criteria for high, medium, and low fidelity as part of the process. Some tools may also have pre-established criteria for implementation fidelity that can be used in the matrix. Stakeholder input in the identification of criteria for fidelity may support greater engagement and buy-in with an initiative, particularly if stakeholders feel as though they have had a role in identifying the criteria by which an initiative is measured. The simple format of this matrix lends itself to facilitated conversations that are broken down by tool and criteria, and the possibility that each tool or dimension of fidelity can be evaluated separately or with unique stakeholders if necessary.

The results matrix allows state teams to quantify and assess activities across multiple levels of implementation into a simplified format. The total overall possible score on this tool is calculated by multiplying the number of tools that inform each dimension of fidelity by three, the maximum number of points assigned for high fidelity of implementation. This study established criteria for high, medium, and low overall implementation fidelity. In order

to achieve an overall rating of high implementation fidelity on this matrix a program would need to receive at least 85% of the total available points. A moderate degree of implementation fidelity would be indicated by an overall score between 60% and 84% and low implementation fidelity would be reflected in an overall score of less than 60% of the total available points. While this criteria is recommended, a state may wish to establish their own criteria in consultation with stakeholders.

Recommendations for Future Research

A number of states are currently implementing or preparing to implement the Pyramid Model statewide (R. Corso, personal communication, December 13, 2016). This study presents information on how one state approached the initial year of implementation in public schools, what implementation science refers to as the initial implementation stage. Researchers may wish to revisit this study and the implementation in Massachusetts to examine the changes in overall statewide implementation fidelity as a result of the modifications made in year two of the initiative and the impact of a multi-year initiative, rather than a single year, on the use of this model by participating districts, including whether or not Massachusetts was able to achieve full implementation. This single case study also reflects the unique challenges and circumstances of Massachusetts' implementation of this model in a state with previous experience with the Pyramid Model. Researchers may wish to conduct a broader study that uses the same results matrix across several different states seeking to scale up implementation of this model in public schools to identify which challenges and opportunities are unique to the Massachusetts context and which are more universal to inform the development of additional resources to support Pyramid Model implementation statewide. Lastly, the timing of this study limits the ability to draw

conclusions about the impact of this implementation on child outcomes, the ultimate purpose of the SSIP. Future research may wish to examine the relationship between implementation fidelity of this model in Massachusetts across multiple years and the outcomes demonstrated by children in classrooms implementing this model as measured by SPP/SPR indicator 7 and/or other child-level outcomes measures.

As described previously, the results matrix has significant potential as a tool for framing the evaluation of, and assessing the implementation fidelity of a statewide initiative. Future research should explore the applicability of this matrix to other states implementing the Pyramid Model and other states implementing different evidence based practices at a statewide scale. While this study utilized this matrix to assess implementation fidelity, the matrix has not been shared with or vetted by MA ESE stakeholders. Future research could examine the tools' use with stakeholders and their input regarding its utility in framing the scope of an evaluation and defining criteria for high implementation fidelity.

Limitations: Threats to external validity.

Selection-treatment interaction. After developing an initial plan for the SSIP, MA ESE began recruiting districts for participation in the spring of 2015 and started working with five districts in May of 2015. An additional 14 districts were recruited over the summer of 2015 and those districts began work under this initiative in September and October of 2015. These 18 districts were identified to participate in the initial Pyramid Model training and implementation based on a number of factors. It was important that participating districts have the capacity for implementation and readiness to take on a substantial initiative. The criteria for inclusion in the initial cohort of districts were established through substantial engagement with stakeholders in order to ensure that these initial participants would be

successful in their implementation. Given the importance of aligning social emotional supports across the grade span, districts that have participated or were scheduled to participate in MA ESE's school-age PBIS trainings were given priority for participating in the Pyramid Model initiative if they assured MA ESE that they were committed to expanding implementation of PBIS into early childhood classrooms through the implementation of the Pyramid Model. In addition, MA ESE gave priority to districts that were participating in MA ESE's Low-income Education Access Project (LEAP) initiative. The LEAP project focuses on providing targeted technical assistance to selected districts to support appropriate identification of students from economic disadvantage in special education and to support their access to inclusive settings once eligible. Through this work, MA ESE had identified the importance of ensuring that students from all backgrounds have the social emotional supports necessary to be successful in their school. MA ESE also prioritized participation of districts receiving an Early Childhood Mental Health System of Care grant because these districts are already committed to undertaking significant work to improve the social emotional development of their students. Lastly, MA ESE worked with other offices in the agency and colleagues in the Department of Early Education and Care (MA EEC) to assess districts' readiness for implementation based on existing knowledge of their current policies and practices, including their current accountability and assistance level.

The decision to recruit districts through a selective outreach process rather than a competitive application may have impacted the initial year of implementation. There is a possibility that direct contact from MA ESE, which oversees district activities, may have led some districts to agree to participate out of a sense of obligation rather than interest. Had the application process been a competitive one, MA ESE may have been more likely to identify

districts who were more prepared to begin implementation immediately or who were already engaged in the initial stages of implementation of the Pyramid Model. Massachusetts has over 400 districts and while these 18 are participating in aligned or similar initiatives, a competitive application process in which districts self-identified might have led to participants who were more interested, excited, and engaged in the project from the outset. In addition the selected districts' participation in other MA ESE initiatives may have contributed to a selection-treatment interaction in which the results of this study are directly related to the relationship between the implementation of the Pyramid Model and the other initiatives these districts were involved in at the time.

Multiple-treatment interference. More than half of the districts participating in this initiative were recruited by MA ESE because of their participation in the MA ESE PBIS Academies. The PBIS Academies are a multi-year training program designed to support the implementation of school-wide PBIS in schools throughout the state. Almost all of the districts participating in both initiatives began their participation in the PBIS academies prior to their participation in the Pyramid Model initiative.

This previous participation in a statewide social emotional tiered intervention framework may have impacted their implementation of the Pyramid Model in early childhood, including the use of age inappropriate strategies for supporting social emotional development from the PBIS trainings in their early childhood classrooms implementing the Pyramid Model. Districts participating in this initiative reported challenges with implementing PBIS in preschool classrooms as many of the strategies utilized in PBIS are not developmentally appropriate for young children (Blair, Fox, & Lentini, 2010). MA ESE worked with its contracted vendors from the University of Connecticut and the Pyramid

Model Consortium to begin to design a more aligned approach between the Pyramid Model and PBIS in Massachusetts. As a result, the Pyramid Model Beginning of the Year Leadership Team Meetings in September and October of 2015 we held on the same dates and at the same locations as the kickoff meetings for schools participating in the PBIS academies and more than half of the districts participating had teams in both the PBIS academy and the Pyramid Model initiative. In addition, experts from the University of Connecticut and the Pyramid Model Consortium co-presented on the opening morning of the first day of the meetings to the preschool Pyramid Model teams and the school age PBIS teams on the alignment between the two models. The intention was that the PBIS and Pyramid Model leadership teams would recognize the importance of aligning their initiatives across the grade span. In addition, ideally this format would have led to PBIS leadership team representation on the Pyramid Model leadership team and vice versa. Participation in both initiatives may have also contributed to strained resources at the local level as individuals may have been asked to participate on multiple leadership teams or attend trainings for both initiatives.

Unfortunately, alignment across the two initiatives was not able to be carried through the rest of the first year of implementation of the Pyramid Model. Scheduling challenges including space availability, presenter availability, and other conflicts led to the overlap between some PBIS and Pyramid Model trainings and activities. In addition, the messaging from the vendors regarding the timeline for implementation was very different. The PBIS Academies are a multi-year project in which the first year is typically spent planning implementation. PBIS implementation schoolwide typically means that a school will have several years to work with the same students to utilize their new practices. However, most preschool programs have students in their classes for a year or two at most, leading to a much

greater sense of urgency in the Pyramid Model trainings regarding the need to implement the Pyramid Model practices as soon as possible.

The prevalence of and recognition for the PBIS Academies as a major statewide initiative may have also contributed to a perception that the Pyramid Model initiative was secondary to or less important than the PBIS academies and lower participant engagement with this initiative. The PBIS Academies began one year prior to the implementation of the Pyramid Model by MA ESE. According to reports from participating districts and the external coaches, the multi-grade, schoolwide emphasis of the PBIS academies almost always led to a prioritization of that initiative at the school and district level as it had the potential to impact more students. This prioritization also meant that preschool teachers were frequently asked to implement strategies such as token systems and office discipline referral tracking that were inappropriate for their context in order to fit into the schoolwide PBIS system. At the Pyramid Model Midyear Leadership Team Meeting the most frequently reported implementation challenge by districts implementing both models was what they perceived as a misalignment between the requirements of participation in the PBIS Academies and what they were being trained to do as part of the Pyramid Model trainings. The external coaches were intended to support the Pyramid Model leadership teams in making connections between the two initiatives and providing resources to leadership teams to help explain that the two models are closely aligned but that the practices may look slightly different due to the differing needs of the student populations they serve. However, the external coaches limited experience with public schools and even more limited experience with schoolwide PBIS implementation meant that their capacity in this area was

very limited. Participation in both initiatives may have also presented challenges for district leadership as they balanced competing priorities from each initiative.

Specificity of variables. The unique circumstances of the implementation of the Pyramid Model in Massachusetts may also contribute to limited external validity for this study. While OSEP required each state to develop and implement a SSIP, the implementation of the Pyramid Model in Massachusetts took place under a unique set of circumstances. Massachusetts is a local control state and had no regulatory authority to mandate implementation of the Pyramid Model in any of the participating districts. Therefore, MA ESE relied on the voluntary participation of districts with limited availability for incentives for participation. In addition, because Massachusetts was a pilot state for the initial testing of the Pyramid Model there was a great deal of familiarity with the Pyramid Model practices among participants in this initiative which in some cases, as indicated by training evaluations, led participants to be less engaged with or excited about training activities as they seemed like a review rather than new content. This preexisting relationship with the developers of the Pyramid Model, including a contract with them to support the implementation of this model in the state, may have also impacted the results of the interviews. The creators, as a contracted vendor and individuals with a vested interest in the success of the model, may have been overly optimistic in their assessment of implementation fidelity as a result of their relationship with MA ESE. This study also utilized a unique MA ESE developed tool, the External Coach Contact Record. Other states may have their own ways in which they assess the frequency, content, and success of coaching activities that may impact the types of and quality of information collected to conduct analyses of implementation fidelity based on the results of that tool.

Lastly, the definitions of high, moderate, and low fidelity presented in Table 7 were selected by the author and reflect the specific Massachusetts context. Another state or entity looking to implement the Pyramid Model may choose to define high, moderate, and low implementation fidelity for each tool differently or may select different tools entirely to inform their analysis of implementation fidelity on each dimension.

Chapter 6

Conclusion

Massachusetts identified poor social emotional outcomes for children with disabilities aged three to five as the focus of its SSIP and selected the Pyramid Model as an evidenced based practice to implement statewide to support improved outcomes for these students. This study presented a process evaluation of the implementation fidelity of this model at the state, district, and classroom level in the approximately first year of implementation from May 2015 to September 2016. Using five different dimensions of implementation fidelity and seven distinct data collection tools, the author created a results matrix which articulates criteria for high, medium, and low implementation fidelity for each tool when it informs a dimension of fidelity. Based on the analyses presented here, Massachusetts achieved a moderate level of implementation fidelity in its first year of this initiative. A number of challenges and opportunities were discussed that can inform other states or large organizations seeking to implement the Pyramid Model. In addition, the results matrix can be customized to reflect the unique tools used by any statewide implementation of an evidenced based practice to simply, concisely, and clearly present the information that will be used to evaluate an initiative and the criteria for success. This is particularly important as more and more states are being asked to undertake large scale initiatives to improve outcomes for students, including those with disabilities and are required to have meaningful stakeholder involvement throughout the process.

References

American Psychological Association. (2008). Are zero tolerance policies effective in the schools?: An evidentiary review and recommendations. *American Psychologist*, *63*(9), 852–862. https://doi.org/10.1037/0003-066X.63.9.852

Anthony, B. J., Anthony, L. G., Morrel, T. M., & Acosta, M. (2005). Evidence for social and behavior problems in low-income, urban preschoolers: Effects of site, classroom, and teacher. *Journal of Youth & Adolescence*, *34*(1), 31–39. https://doi.org/10.1007/s10964-005-1334-y

Aro, T., Laakso, M.-L., Määttä, S., Tolvanen, A., & Poikkeus, A.-M. (2014). Associations between toddler-age communication and kindergarten-age self-regulatory skills. *Journal of Speech, Language & Hearing Research*, *57*(4), 1405–1417. https://doi.org/10.1044/2014

Baker, B. L., McIntyre, L. L., Blacher, J., Crnic, K., Edelbrock, C., & Low, C. (2003). Pre-school children with and without developmental delay: Behaviour problems and parenting stress over time. *Journal of Intellectual Disability Research*, *47*. Retrieved from http://escholarship.org/uc/item/5f28z26b

Barber, A. B., Saffo, R. W., Gilpin, A. T., Craft, L. D., & Goldstein, H. (2016). Peers as clinicians: Examining the impact of Stay Play Talk on social communication in young preschoolers with autism. *Journal of Communication Disorders*, *59*, 1–15. https://doi.org/10.1016/j.jcomdis.2015.06.009

Belinda Ratcliffe, M. W. (2014). Teaching social–emotional skills to school-aged children with Autism Spectrum Disorder: A treatment versus control trial in 41 mainstream

schools. *Research in Autism Spectrum Disorders*, *8*(12), 1722–1733. https://doi.org/10.1016/j.rasd.2014.09.010

Benedict, E. A., Horner, R. H., & Squires, J. K. (2007). Assessment and implementation of positive behavior support in preschools. *Topics in Early Childhood Special Education*, *27*(3), 174–192. https://doi.org/10.1177/02711214070270030801

Bennett, P., Elliott, M., & Peters, D. (2005). Classroom and family effects on children's social and behavioral problems. *Elementary School Journal*, *105*(5), 461–480.

Bertram, R. M., Blase, K. A., & Fixsen, D. L. (2015). Improving programs and outcomes: Implementation frameworks and organization change. *Research on Social Work Practice*, *25*(4), 477–487. https://doi.org/10.1177/1049731514537687

Bethell, C. D., Read, D., Blumberg, S. J., & Newacheck, P. W. (2007). What is the prevalence of children with special health care needs? Toward an understanding of variations in findings and methods across three national surveys. *Maternal and Child Health Journal*, *12*(1), 1–14. https://doi.org/10.1007/s10995-007-0220-5

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, *57*(2), 111.

Blair, K.-S. C. (2010). Use of positive behavior support to address the challenging behavior of young children within a community early childhood program. *Topics in Early Childhood Special Education*, *30*(2), 68–79. https://doi.org/10.1177/0271121410372676

Blair, K.-S. C., Fox, L., & Lentini, R. (2010). Use of positive behavior support to address the challenging behavior of young children within a community early childhood program. *Topics in Early Childhood Special Education*, *30*(2), 68–79.

Blase, K., Kiser, L. and Van Dyke, M. (2013). *The Hexagon Tool: Exploring Context.* Chapel Hill, NC: National Implementation Research Network, FPG Child Development Institute, University of North Carolina at Chapel Hill.

Bradshaw, C. P., Reinke, W. M., Brown, L. D., Bevans, K. B., & Leaf, P. J. (2008). Implementation of school-wide positive behavioral interventions and supports (PBIS) in elementary schools: Observations from a randomized trial. *Education and Treatment of Children*, *31*(1), 1–26.

Brown, W., H., & Conroy, M., A. (2011). Social-emotional competence in young children with developmental delays: our reflection and vision for the future. *Journal of Early Intervention*, *33*(4), 310–320. https://doi.org/10.1177/1053815111429969

Bulotsky-Shearer, R. J., & Fantuzzo, J. W. (2011). Preschool behavior problems in classroom learning situations and literacy outcomes in kindergarten and first grade. *Early Childhood Research Quarterly*, *26*(1), 61–73. https://doi.org/10.1016/j.ecresq.2010.04.004

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction. *Applied Development Science*, *12*(3), 140–153.

Campbell, P. H., & Anketell, M. (2007). Suggestions for statewide measurement systems Pennsylvania's experience. *Topics in Early Childhood Special Education*, *27*(1), 34–48. https://doi.org/10.1177/02711214070270010301

Cannon, Y., Gregory, M., & Waterstone, J. (2013). A solution hiding in plain sight: Special education and better outcomes for students with social, emotional, and behavioral challenges. *Fordham Urban Law Journal*, *41*(2), 403–497.

Carter, D. R., & Van Norman, R. K. (2010). Class-wide positive behavior support in preschool: Improving teacher implementation through consultation. *Early Childhood Education Journal*, *38*(4), 279–288. https://doi.org/ 10.1007/s10643-010-0409-x

Carter, D. R., Van Norman, R. K., & Tredwell, C. (2011). Program-wide positive behavior support in preschool: Lessons for getting started. *Early Childhood Education Journal*, *38*(5), 349–355. https://doi.org/10.1007/s10643-010-0406-0

Cheng, E. R., Palta, M., Kotelchuck, M., Poehlmann, J., & Witt, W. P. (2014). Cognitive delay and behavior problems prior to school age. *Pediatrics*, *134*(3), e749–e757. https://doi.org/10.1542/peds.2014-0259

Cheng, Park, H., Robert, S. A., Palta, M., & Witt, W. P. (2014). Impact of county disadvantage on behavior problems among US children with cognitive delay. *American Journal of Public Health*, *104*(11), 2114–2121. https://doi.org/10.2105/AJPH.2014.302119

Codding, R. S., Skowron, J., & Pace, G. M. (2005). Back to basics: training teachers to interpret curriculum-based measurement data and create observable and measurable objectives. *Behavioral Interventions*, *20*(3), 165–176. https://doi.org/10.1002/bin.194

Cohen, D. K., & Hill, H. C. (2008). *Learning Policy: When State Education Reform Works*. Yale University Press.

Crnic, K. A., Neece, C. L., McIntyre, L. L., Blacher, J., & Baker, B. L. (2017). Intellectual disability and developmental risk: Promoting intervention to improve child and family well-being. *Child Development*, n.a., 1-10. https://doi.org/10.1111/cdev.12740

Crnic, K., Hoffman, C., Gaze, C., & Edelbrock, C. (2004). Understanding the emergence of behavior problems in young children with developmental delays. *Infants & Young Children*, *17*(3), 223–235.

CSEFEL. (2015). *The Center for the Social Emotional Foundations of Early Learning*. Retrieved from http://csefel.vanderbilt.edu/

Davis, E., Corr, L., Ummer-Christian, R., Gilson, K.-M., Waters, E., Mihalopoulos, C., … Sims, M. (2014). Family day care educator's Knowledge, confidence and skills in promoting children's's social and emotional wellbeing: Baseline data from Thrive. *Australasian Journal of Early Childhood*, *39*(3), 66–75.

Domínguez, X., Vitiello, V. E., Fuccillo, J. M., Greenfield, D. B., & Bulotsky-Shearer, R. (2011). The role of context in preschool learning: A multilevel examination of the contribution of context-specific problem behaviors and classroom process quality to low-income children's approaches to learning. *Journal of School Psychology*, *49*(2), 175–195. https://doi.org/10.1016/j.jsp.2010.11.002

Duda, M. A., Dunlap, G., Fox, L., Lentini, R., & Clarke, S. (2004). An Experimental evaluation of positive behavior support in a community preschool program. *Topics in Early Childhood Special Education*, *24*(3), 143–155.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research*, *18*(2), 237–256. https://doi.org/10.1093/her/18.2.237

Early, D. M., Maxwell, K. L., Burchinal, M., Bender, R. H., Ebanks, C., Henry, G. T., … Vandergrift, N. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. *Child Development*, *78*(2), 558–580. https://doi.org/10.1111/j.1467-8624.2007.01014.x

ECTA. (2015). *The Early Childhood Technical Assistance Center*. Retrieved from http://ectacenter.org/eco/pages/faqs.asp#Overview

ECTA Center. (2015, May). *Office of Special Education Program SSIP Phase I Implementation Guide - Part B.* Retrieved from http://ectacenter.org/~pdfs/topics/ssip/Part-B-Implementation-Guide-May7.pdf

Emerson, E., & Einfeld, S. (2010). Emotional and behavioural difficulties in young children with and without developmental delay: a bi-national perspective. *Journal of Child Psychology & Psychiatry*, *51*(5), 583–593. https://doi.org/10.1111/j.1469-7610.2009.02179.x

Emerson, E., & Hatton, C. (2007). Poverty, socio-economic position, social capital and the health of children and adolescents with intellectual disabilities in Britain: a replication. *Journal of Intellectual Disability Research*, *51*(11), 866–874. https://doi.org/10.1111/j.1365-2788.2007.00951.x

OSEP. (2006). Monitoring, technical assistance and enforcement. Retrieved from http://idea.ed.gov/explore/view/p/%2Croot%2Cdynamic%2CTopicalBrief%2C24%2C

OSEP. (2013). Part B SPP/APR related requirements. Retrieved from http://www2.ed.gov/policy/speced/guid/idea/bapr/2014/2014brelatedrequirements.pdf

Fixsen, D., Blase, K., Metz, A., & Dyke, M. V. (2013). Statewide implementation of evidence-based programs. *Exceptional Children*, *79*(2), 213–230.

Fox, L., Hemmeter, M., Snyder, P., Binder, D. P., & Clarke, S. (2011). Coaching early childhood special educators to implement a comprehensive model for promoting young children's social competence. *Topics in Early Childhood Special Education*, *31*(3), 178–192. https://doi.org/10.1177/0271121411404440

Fox, L., & Smith, B. J. (2007). Promoting social, emotional and behavioral outcomes of young children served under IDEA. Issue Brief. *Technical Assistance Center on Social Emotional Intervention for Young Children*. Retrieved from http://eric.ed.gov/?id=ED526382

Frey A. J., Small J. W., Lee J., Walker H. M., Seeley J. R., Feil E. G., Golly A. (2015). Expanding the range of the First Step to Success intervention: Tertiary-level support for children, teachers, and families. *Early Childhood Research Quarterly*, *30*, 1–11.

Gebbie, D., Ceglowski, D., Taylor, L., & Miels, J. (2012). The Role of teacher efficacy in strengthening classroom support for preschool children with disabilities who exhibit challenging behaviors. *Early Childhood Education Journal*, *40*(1), 35–46. https://doi.org/10.1007/s10643-011-0486-5

Geldart, S. (2016, March). *Preschool PBS through pyramid strategies mid-year leadership team meeting*. Presentation presented at the Preschool PBS through Pyramid Strategies Mid-Year Leadership Team Meeting, Marlborough, MA.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*(2), 149–164. https://doi.org/10.1177/001440290507100202

Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems*. Foundation for Child Development. Retrieved from http://www.childstudycenter.yale.edu/zigler/publications/34774_National%20Prek%20Study_expulsion.pdf

Green, S. E. (2007). "We're tired, not sad": Benefits and burdens of mothering a child with a disability. *Social Science & Medicine*, *64*(1), 150–163. https://doi.org/10.1016/j.socscimed.2006.08.025

Greenwood, C. R., Walker, D., Hornbeck, M., Hebbeler, K., & Spiker, D. (2007). Progress developing the Kansas early childhood special education accountability system: Initial findings using ECO and COSF. *Topics in Early Childhood Special Education*, *27*(1), 2–18.

Hall, T. P. , Turnbull, A. P. , McCart, A., Griggs, P., Jeong-Hoon, C. , Markey, U., … Sailor, W. (2007). The Effects of positive behavior support parent-training programs on parent-child relationships in culturally and linguistically diverse families. *Multiple Voices for Ethnically Diverse Exceptional Learners*, 10(1/2), 191–210.

Hebbeler, K., & Rooney, R. (2009). Accountability for services for young children with disabilities and the assessment of meaningful outcomes: The role of the speech-language pathologist. *Language, Speech, and Hearing Services in Schools*, *40*(4), 446–456. https://doi.org/10.1044/0161-1461(2009/08-0025)

Hemmeter, M. L., Santos, R. M., & Ostrosky, M. M. (2008). Preparing early childhood educators to address young children's social-emotional development and challenging behavior: a Survey of higher education programs in nine states. *Journal of Early Intervention*, *30*(4), 321–340.

Hemmeter, M. L., Snyder, P. A., Fox, L., & Algina, J. (2016). Evaluating the implementation of the Pyramid Model for Promoting Social-Emotional Competence in Early Childhood Classrooms. *Topics in Early Childhood Special Education*, *36*(3), 133–146. https://doi.org/10.1177/0271121416653386

Hemmeter, M. L., Fox, L., & Snyder, P. (2013). *Teaching Pyramid Observation Tool (TPOT) for Preschool Classrooms Manual.* Brookes Publishing.

Hoch, J., Spofford, L., Dimian, A., Tervo, R., MacLean, W. E., & Symons, F. J. (2016). A Direct comparison of self-injurious and stereotyped motor behavior between preschool-aged children with and without developmental delays. *Journal of Pediatric Psychology*, *41*(5), 566–572. https://doi.org/10.1093/jpepsy/jsv102

Horner, R. H., Carr, E. G., Strain, P. S., Todd, A. W., & Reed, H. K. (2002). Problem behavior interventions for young children with autism: A Research synthesis. *Journal of Autism and Developmental Disorders*, *32*(5), 423–446. https://doi.org/10.1023/A:1020593922901

Individuals With Disabilities Education Act, 20 U.S.C. § 1400 (2004).

Johnston, J. , Foxx, R. M., Jacobson, J. W., Green, G., & Mulick, J. A. (2006). Positive behavior support and applied behavior analysis. *The Behavior Analyst*, *29*(1), 51–74.

Kasprzak, C., Hurth, J., Rooney, R., Goode, S. E., Danaher, J. C., Whaley, K. T., … Cate, D. (2011). States' accountability and progress in serving young children with disabilities.

Topics in Early Childhood Special Education, 27. https://doi.org/10.1177/0271121411408119

Kathleen J. Marshall, William H. Brown, Maureen A. Conroy, & Herman Knopf. (2011). Early intervention and prevention of disability. In *Handbook of Special Education* (pp. 703–715). New York, NY.

Klingner, J. K., Boardman, A. G., & McMaster, K. L. (2013). What does it take to scale up and sustain evidence-based practices? *Exceptional Children*, *79*(2), 195–211. https://doi.org/10.1177/001440291307900205

Kohen, D. E., Leventhal, T., Dahinten, V. S., & McIntosh, C. N. (2008). Neighborhood disadvantage: Pathways of effects for young children. *Child Development*, *79*(1), 156–169. https://doi.org/10.1111/j.1467-8624.2007.01117.x

Lamont, J. H., Devore, C. D., Allison, M., Ancona, R., Barnett, S. E., Gunther, R., … Young, T. (2013). Out-of-school suspension and expulsion. *Pediatrics*, *131*(3), e1000–e1007. https://doi.org/10.1542/peds.2012-3932

Lee, K., Calkins, A., & Shin, T. S. (2016). Head Start's impact on social–emotional outcomes for children With disabilities. *Research on Social Work Practice*, *26*(7), 790–802. https://doi.org/10.1177/1049731514568024

Losen, D. J. (2014). *Closing the School Discipline Gap: Equitable Remedies for Excessive Exclusion*. Teachers College Press.

Ma, X., Shen, J., Krenn, H. Y., Hu, S., & Yuan, J. (2016). A Meta-analysis of the relationship between learning outcomes and parental involvement during early childhood education and early elementary education. *Educational Psychology Review*, *28*(4), 771–801. https://doi.org/10.1007/s10648-015-9351-1

Magnuson, K. A., & Waldfogel, J. (2005). Early childhood care and education: Effects on ethnic and racial gaps in school readiness. *The Future of Children*, *15*(1), 169–196. https://doi.org/10.1353/foc.2005.0005

Markowitz, J., Carlson, E., Frey, W., Riley, J., Shimshak, A., Heinzen, H., … Klein, S. (2006). *Preschoolers with disabilities: Characteristics, services, and results. Wave 1 overview report from the pre-elementary education longitudinal study (PEELS).* National Center for Special Education Research.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., … Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–749. https://doi.org/10.1111/j.1467-8624.2008.01154.x

MA ESE. (2014a). *ESE strategic plan.* Retrieved from http://www.doe.mass.edu/research/StrategicPlan.pdf

MA ESE. (2014b). *Massachusetts part B state annual performance report (MA APR) for FFY2012*. Malden, MA.

MA ESE. (2015a). *Enrollment of students with disabilities.* Retrieved from http://www.doe.mass.edu/infoservices/reports/enroll/

MA ESE. (2015b). *Massachusetts part B state annual performance report (MA APR) for FFY2013*. Malden, MA.

MA ESE (2016). *Massachusetts part B state annual performance report (MA APR) for FFY2014*. Malden, MA.

McCabe, P. C., & Meller, P. J. (2004). The relationship between language and social competence: How language impairment affects social growth. *Psychology in the Schools*, *41*(3), 313–321. https://doi.org/10.1002/pits.10161

Muscott, H. S., Mann, E. L., & LeBrun, M. R. (2008). Positive behavioral interventions and supports in New Hampshire: Effects of large-scale implementation of schoolwide positive behavior support on student discipline and academic achievement. *Journal of Positive Behavior Interventions*, *10*(3), 190–205.

Muscott, H. S., Pomerleau, T., & Szczesiul, S. (2009). Large-scale implementation of program-wide positive behavioral interventions and supports in early childhood education programs in New Hampshire. *NHSA DIALOG*, *12*(2), 148–169.

Musgrove, M. (2015, June 30*). Massachusetts part B determination letter*. Retrieved from http://www2.ed.gov/fund/data/report/idea/partbspap/2015/ma-acc-aprltr-2015b.pdf

Nagro, S. A., & Cornelius, K. E. (2013). Evaluating the evidence base of video analysis: A special education teacher development tool. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *36*(4), 312–329. https://doi.org/10.1177/0888406413501090

National Survey of Early Care and Education. (2013). *Number and characteristics of early care and education (ECE) teachers and caregivers: Initial findings from the National Survey of Early Care and Education (NSECE)*. Washington, D.C.

Nelson, C. A., & Luciana, M. (2001). *Handbook of Developmental Cognitive Neuroscience*. MIT Press.

NIRN (2017). *Stages of Implementation Analysis: Where Are We?* Retrieved from: http://implementation.fpg.unc.edu/resources/stages-implementation-analysis-where-are-we?o=nirn

Odom, S. L. (2009). The tie that binds: Evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education*, *29*(1), 53–61. https://doi.org/10.1177/0271121408329171

O'Donnell, C. L. (2008a). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention Research. *Review of Educational Research*, *78*(1), 33–84. https://doi.org/10.3102/0034654307313793

Ogden, T., & Fixsen, D. L. (2014). Implementation science: A brief overview and a look ahead. *Zeitschrift Für Psychologie*, *222*(1), 4–11. https://doi.org/10.1027/2151-2604/a000160

OSEP. (2014). Part B state systemic improvement plan (SSIP) OSEP evaluation tool. Federal Office of Special Education Programs. Retrieved from httpshttps://osep.grads360.org/#communities/pdc/documents/6454://osep.grads360.org/#

OSEP. (2015). *Results driven accountability*. Retrieved from http://www2.ed.gov/about/offices/list/osers/osep/rda/index.html

Petras, H., Masyn, K. E., Buckley, J. A., Ialongo, N. S., & Kellam, S. (2011). Who is most at risk for school removal? A multilevel discrete-time survival analysis of individual- and context-level influences. *Journal of Educational Psychology*, *103*(1), 223–237. https://doi.org/10.1037/a0021545

Poulu, M. (n.d.). Emotional and behavioural difficulties in preschool. *Journal of Child and Family Studies*, *24*, 225–236.

Qi, C. H., & Kaiser, A. P. (2003). Behavior problems of preschool children from low-income families review of the literature. *Topics in Early Childhood Special Education*, *23*(4), 188–216. https://doi.org/10.1177/02711214030230040201

Raffaele Mendez, L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development*, *2003*(99), 17–33. https://doi.org/10.1002/yd.52

Raver, C. C., Blair, C., & Garrett-Peters, P. (2014). Poverty, household chaos, and interparental aggression predict children's ability to recognize and modulate negative emotions. *Development and Psychopathology*, *FirstView*, 1–14. https://doi.org/10.1017/S0954579414000935

Reardon, S. F., & Portilla, X. A. (2016). Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry. *AERA Open*, *2*(3), 2332858416657343. https://doi.org/10.1177/2332858416657343

Rimm-Kaufman, S. E., Pianta, R. C., & Cox, M. J. (2000). Teachers' judgments of problems in the transition to kindergarten. *Early Childhood Research Quarterly*, *15*(2), 147–166. https://doi.org/10.1016/S0885-2006(00)00049-1

Ryder, R. (2016). *Massachusetts part B determination letter*. Retrieved from http://www2.ed.gov/fund/data/report/idea/partbspap/2016/ma-aprltr-2016b.pdf

Saddler, H. (2014). Researching the influence of teaching assistants on the learning of pupils identified with special educational needs in mainstream primary schools: Exploring social inclusion. *Journal of Research in Special Educational Needs*, *14*(3), 145–152.

Sandall, S. R., Schwartz, I. S., & Lacroix, B. (2004). Interventionists' perspectives about data collection in integrated early childhood classrooms. *Journal of Early Intervention*, *26*(3), 161–174. https://doi.org/10.1177/105381510402600301

Schepis, M. M., Ownbey, J. B., Parsons, M. B., & Reid, D. H. (2000). Training support staff for teaching young children with disabilities in an inclusive preschool setting. *Journal of Positive Behavior Interventions*, *2*(3), 170–178. https://doi.org/10.1177/109830070000200305

Schmitt, S. A., Pratt, M. E., & McClelland, M. M. (2014). Examining the validity of behavioral self-regulation tools in predicting preschoolers' academic achievement. *Early Education and Development*, *25*(5), 641–660. https://doi.org/10.1080/10409289.2014.850397

Schoen, J. P., & John W. (2017). Trump's education nominee DeVos confirmed as Pence casts historic tie-breaking vote. Retrieved from http://www.cnbc.com/2017/02/07/trumps-education-nominee-devos-confirmed-as-pence-casts-historic-tiebreaking-vote.html

Simon, A. E., Pastor, P. N., Avila, R. M., & Blumberg, S. J. (2013). Socioeconomic disadvantage and developmental delay among US children aged 18 months to 5 years. *Journal of Epidemiology and Community Health*, jech-2013-202610. https://doi.org/10.1136/jech-2013-202610

Snyder, P. A., Hemmeter, M. L., Fox, L., Bishop, C. C., & Miller, M. D. (2013). Developing and gathering psychometric evidence for a fidelity instrument: The teaching pyramid observation tool–pilot version. *Journal of Early Intervention*, 1053815113516794. https://doi.org/10.1177/1053815113516794

Stoiber, K. C., & Gettinger, M. (2011). Functional assessment and positive support strategies for promoting resilience: Effects on teachers and high-risk children. *Psychology in the Schools*, *48*(7), 686–706.

The Pyramid Model Consortium: About Us. (2016). Retrieved from http://www.pyramidmodel.org/about-us.html

The White House, Office of the Press Secretary. (2014). Fact sheet: Invest in US: The white house summit on early childhood education. Retrieved from http://www.whitehouse.gov/the-press-office/2014/12/10/fact-sheet-invest-us-white-house-summit-early-childhood-education

U.S. Department of Education Office for Civil Rights. (2014). *Data snapshot: Early childhood education*. Washington, D.C.: U.S. Department of Education.

U.S. Department of Health and Human Services, U. S. D. of E. (2014). *Policy statement on expulsion and suspension policies in early childhood settings*. Washington, D.C.: U.S. Department of Health and Human Services.

Vitiello, V. E., Booren, L. M., Downer, J. T., & Williford, A. P. (2012). Variation in children's classroom engagement throughout a day in preschool: Relations to classroom and child factors. *Early Childhood Research Quarterly*, *27*(2), 210–220. https://doi.org/10.1016/j.ecresq.2011.08.005

Voorhees, M. D., Walker, V. L., Snell, M. E., & Smith, C. G. (2013). A demonstration of individualized positive behavior support interventions by Head Start staff to address children's challenging behavior. *Research & Practice for Persons with Severe Disabilities*, *38*(3), 173–185. https://doi.org/10.1177/154079691303800304

Williamson, L., Davis, E., Priest, N., & Harrison, L. (2011). Australian family day care educators: A snapshot of their qualifications, training and perceived support. *Australasian Journal of Early Childhood*, *36*(4), 63.

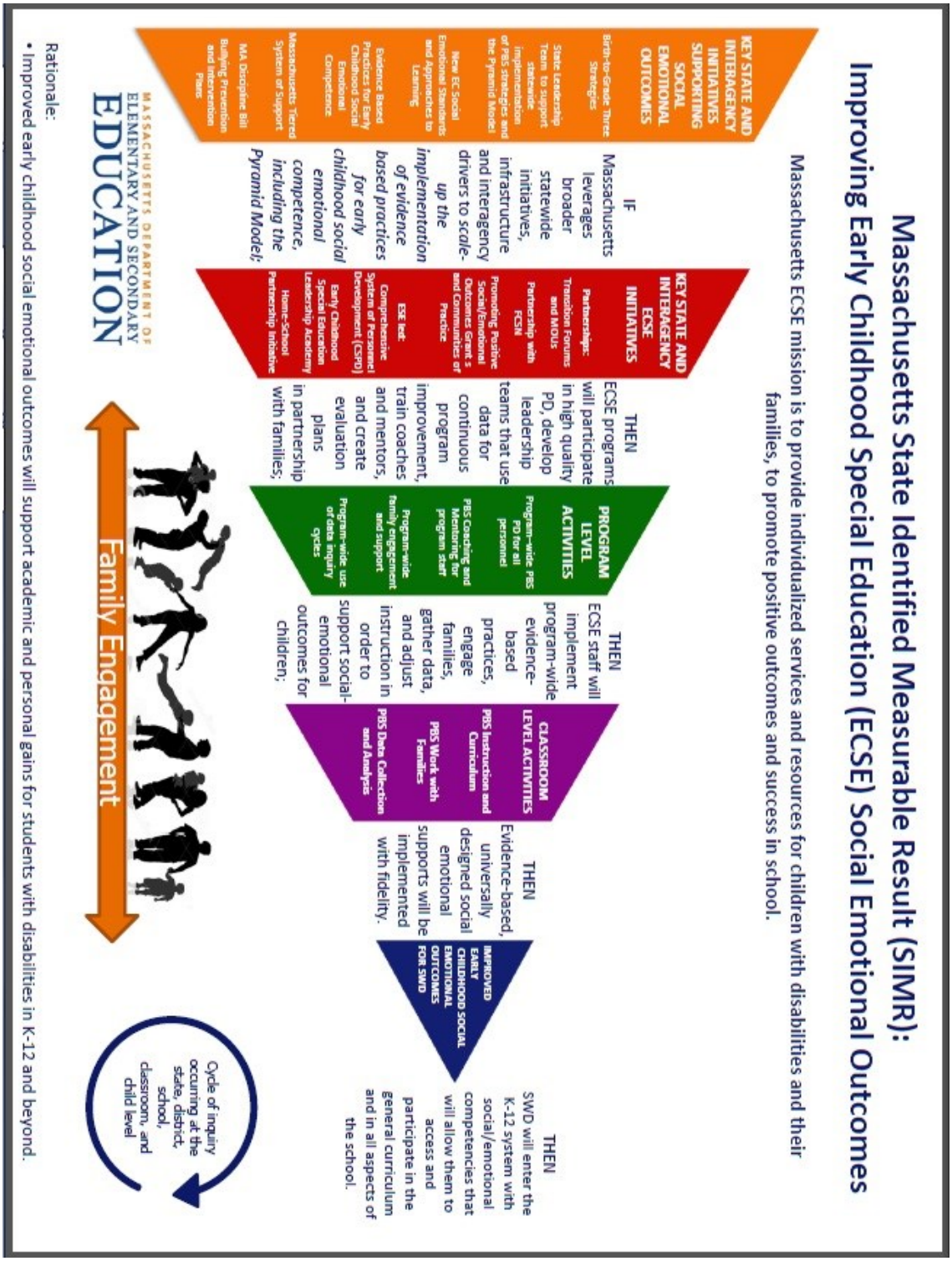


Figure 1. Theory of Change for Improving Social Emotional Outcomes for Young Children with Disabilities in Massachusetts.

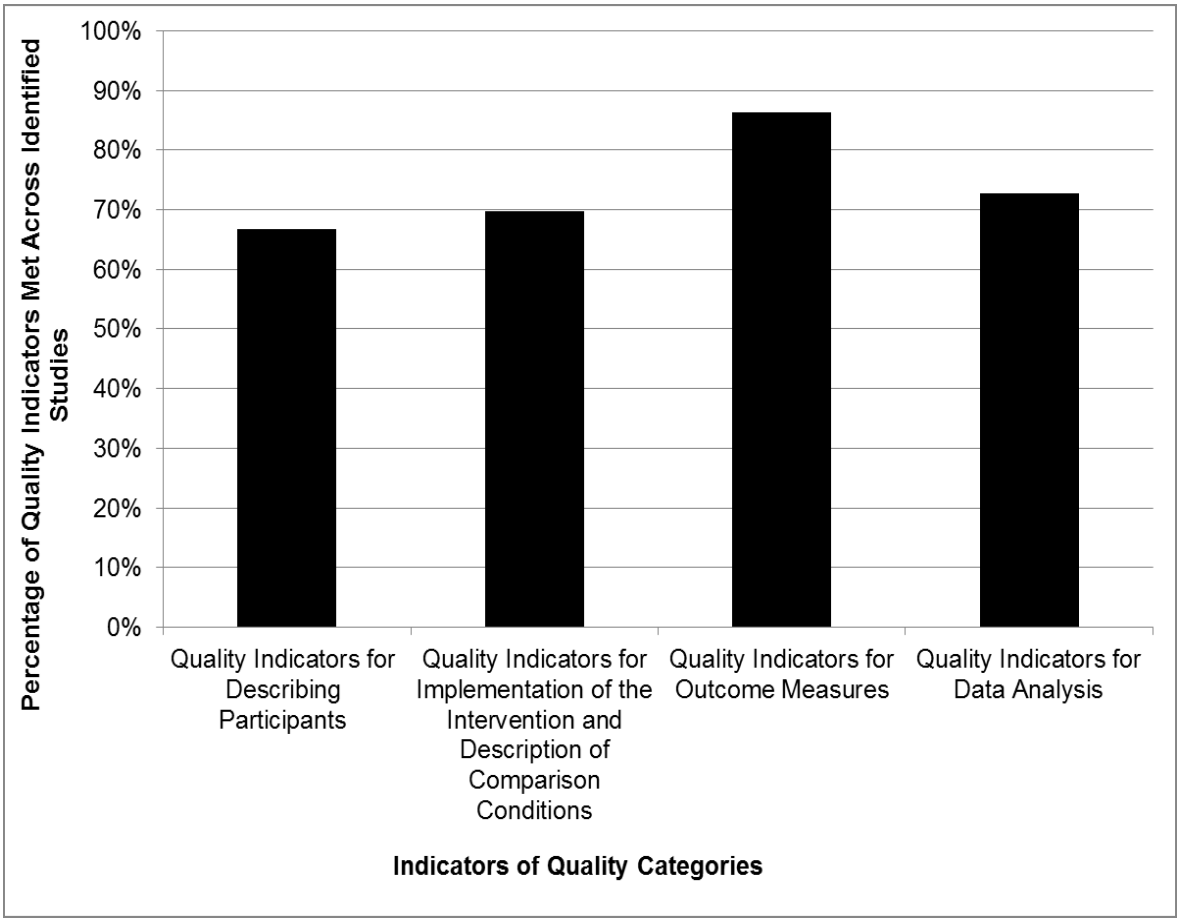
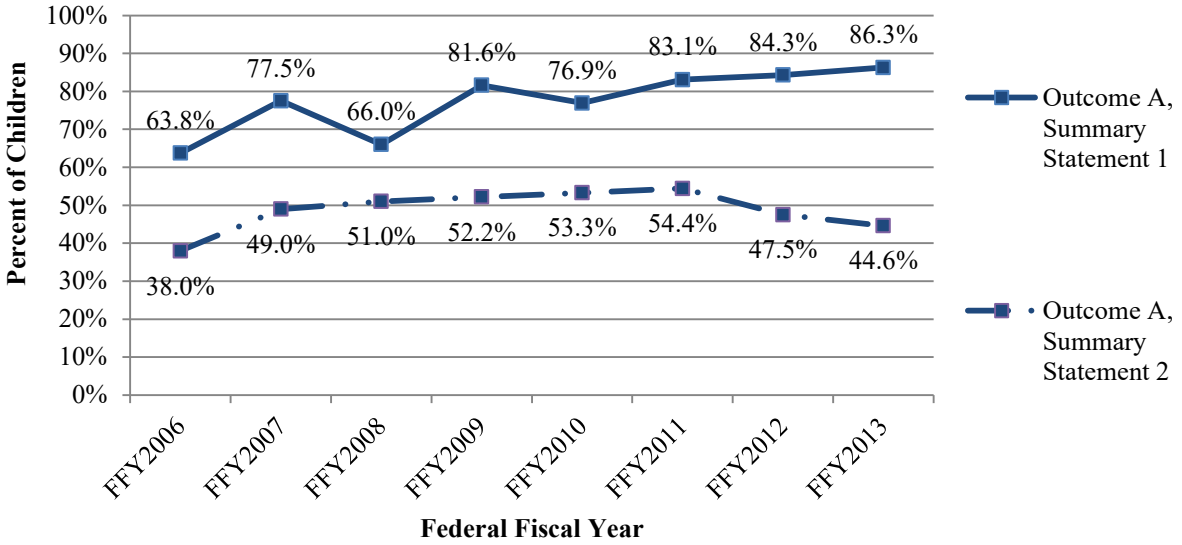


Figure 2. Grand mean percentages on indicators of quality across studies. This figure presents the results of an analysis of the grand mean percentages on the four high quality research domains as established by Gersten et al. (2005).

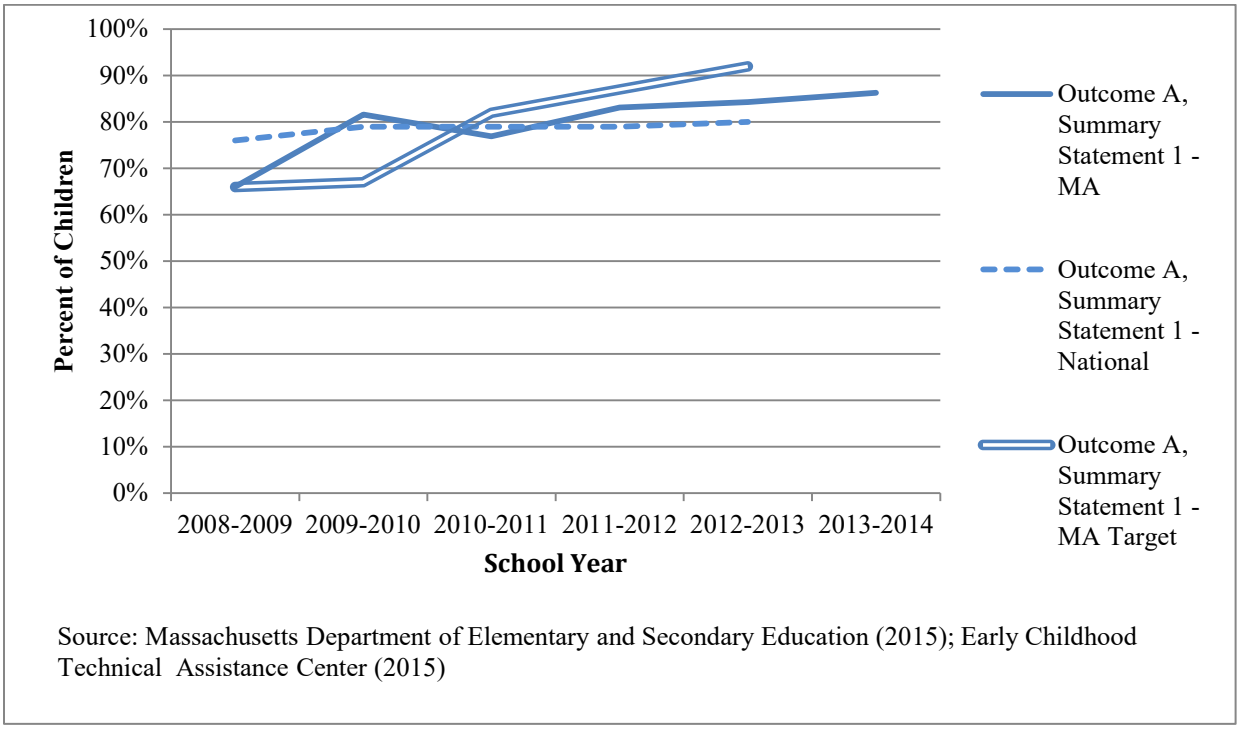


Summary Statement 1: Of those children who entered or exited below age expectations in each outcome, the percent who substantially increased their rate of growth by exit.
Summary Statement 2: The percent of children who were functioning within age expectations by the time they exited

Source: Massachusetts Department of Elementary and Secondary Education (2015).

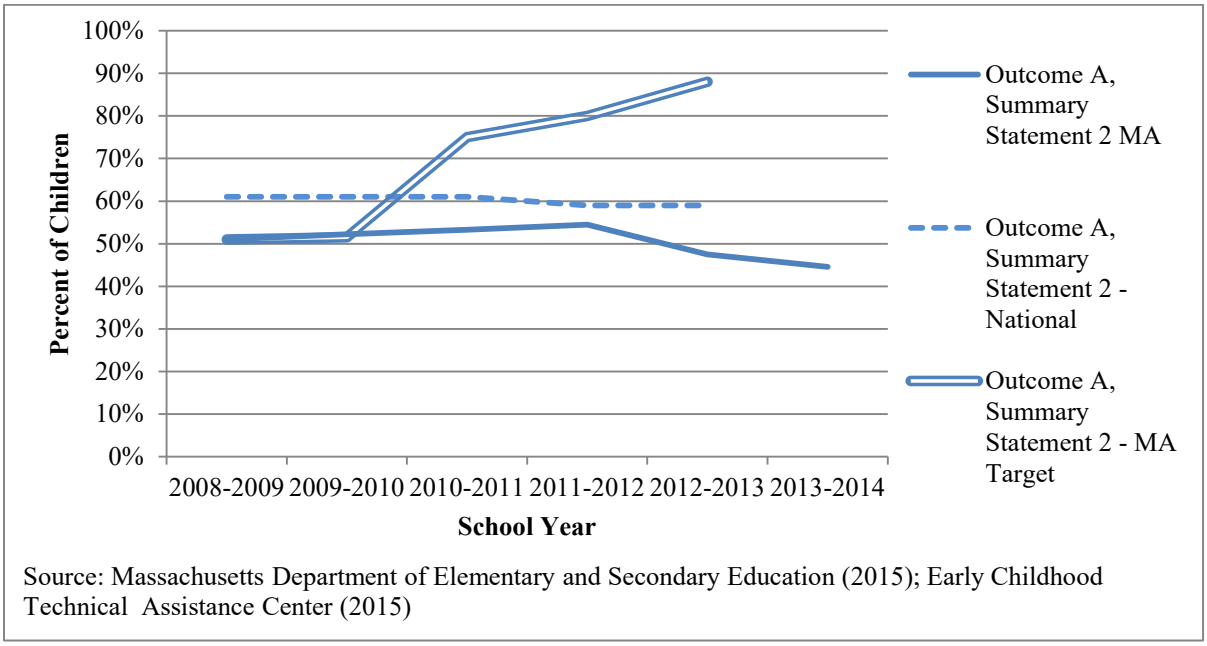
Figure 3. Longitudinal summary statement values for Massachusetts beginning in school year

2006-2007 (Federal Fiscal Year (FFY) 2006) for Outcome A.



Source: Massachusetts Department of Elementary and Secondary Education (2015); Early Childhood Technical Assistance Center (2015)

Figure 4. Longitudinal Comparison of the National Data, MA Data, and MA Targets for Outcome A (Social Emotional Skills) Summary Statement 1.



Source: Massachusetts Department of Elementary and Secondary Education (2015); Early Childhood Technical Assistance Center (2015)

Figure 5. Longitudinal Comparison of the National Data, MA Data, and MA Targets for Outcome A (Social Emotional Skills) Summary Statement 2.

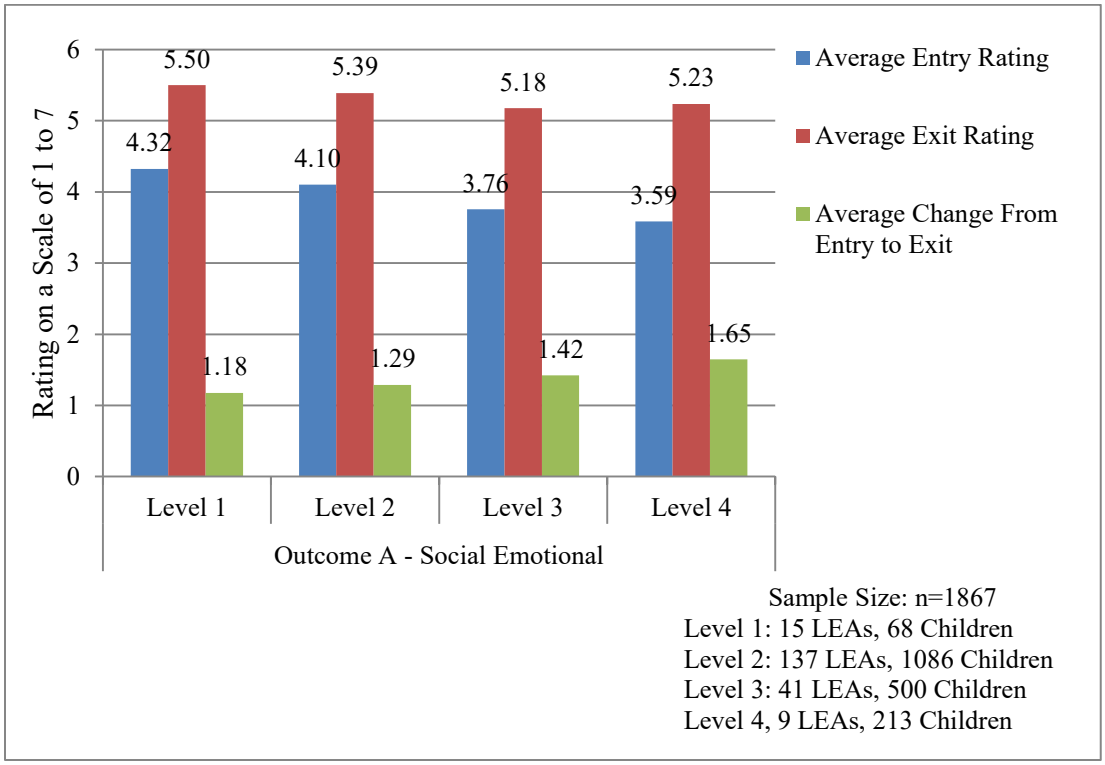


Figure 6. Summary statement values for the social emotional outcome area by district accountability and assistance level.

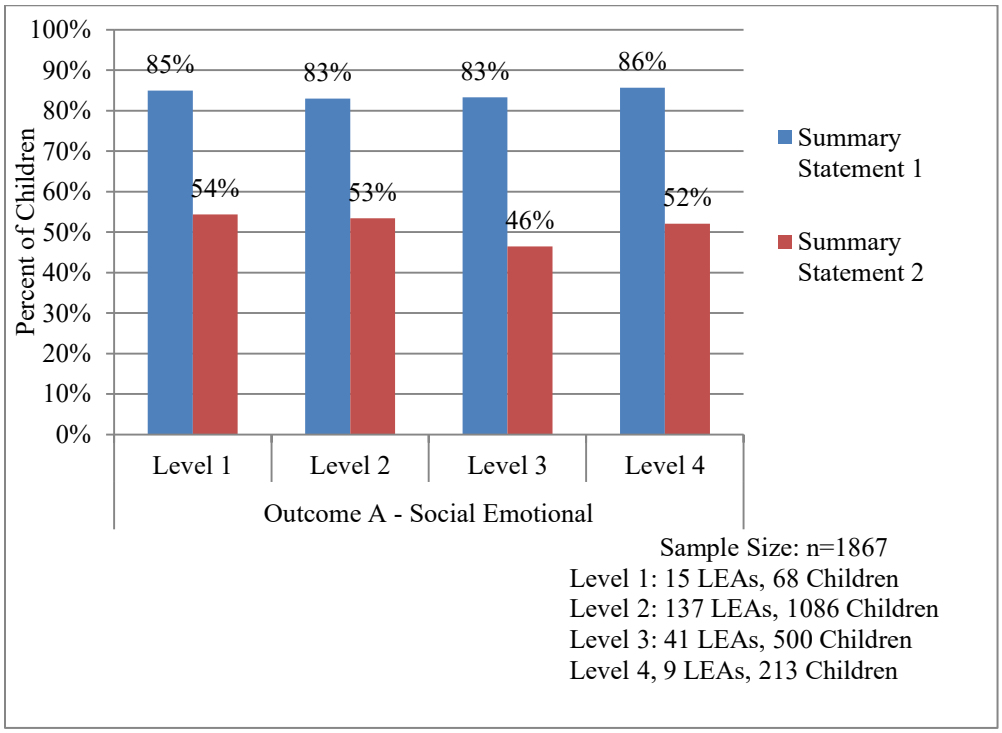


Figure 7. Summary Statement Values by District Accountability and Assistance Level 2012 & 2013 Exiting Children.

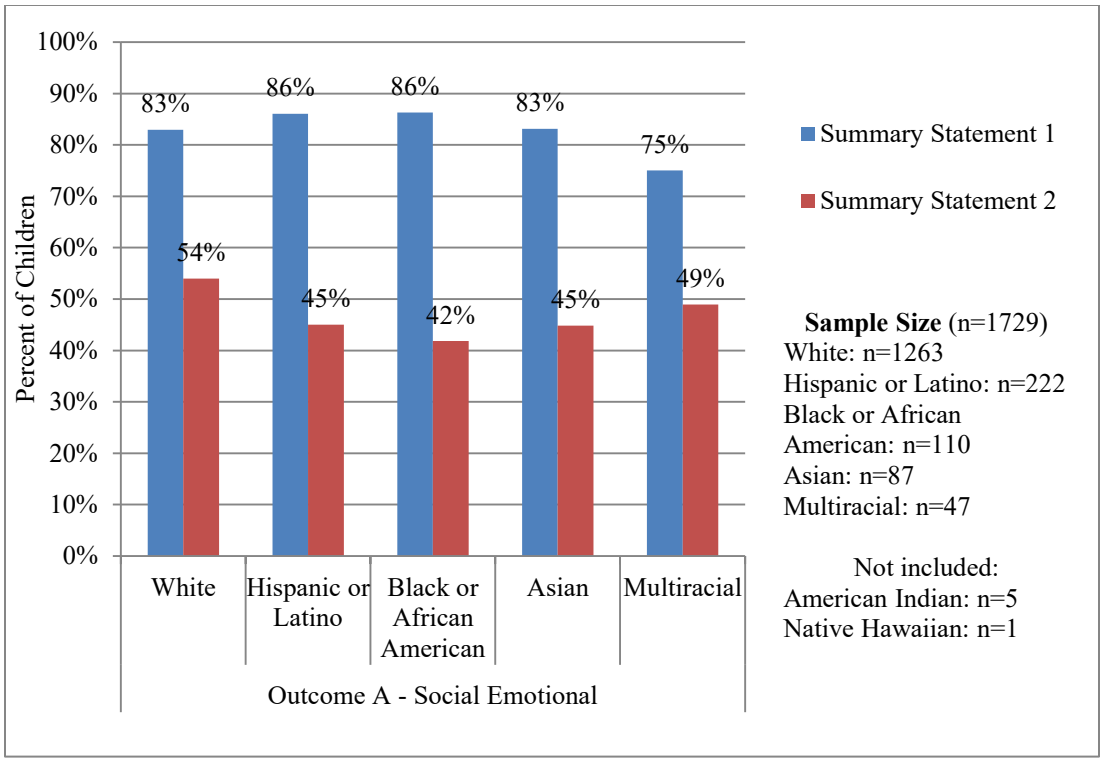


Figure 8. Summary Statement Values by Student Race, 2012 & 2013 Exiting Children.

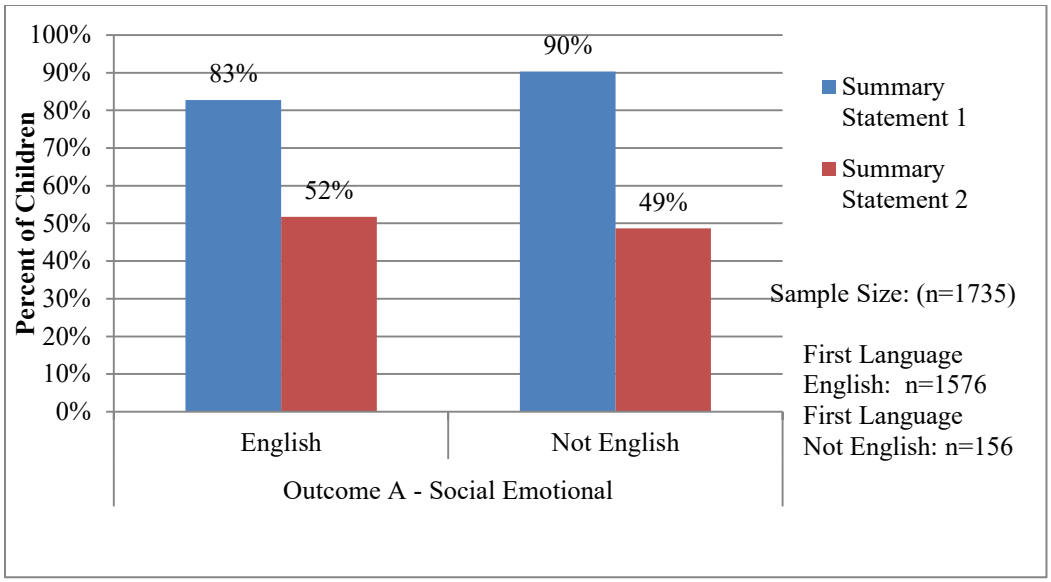


Figure 9. Summary Statement Values by First Language Not English Status, 2012 & 2013 Exiting Children.

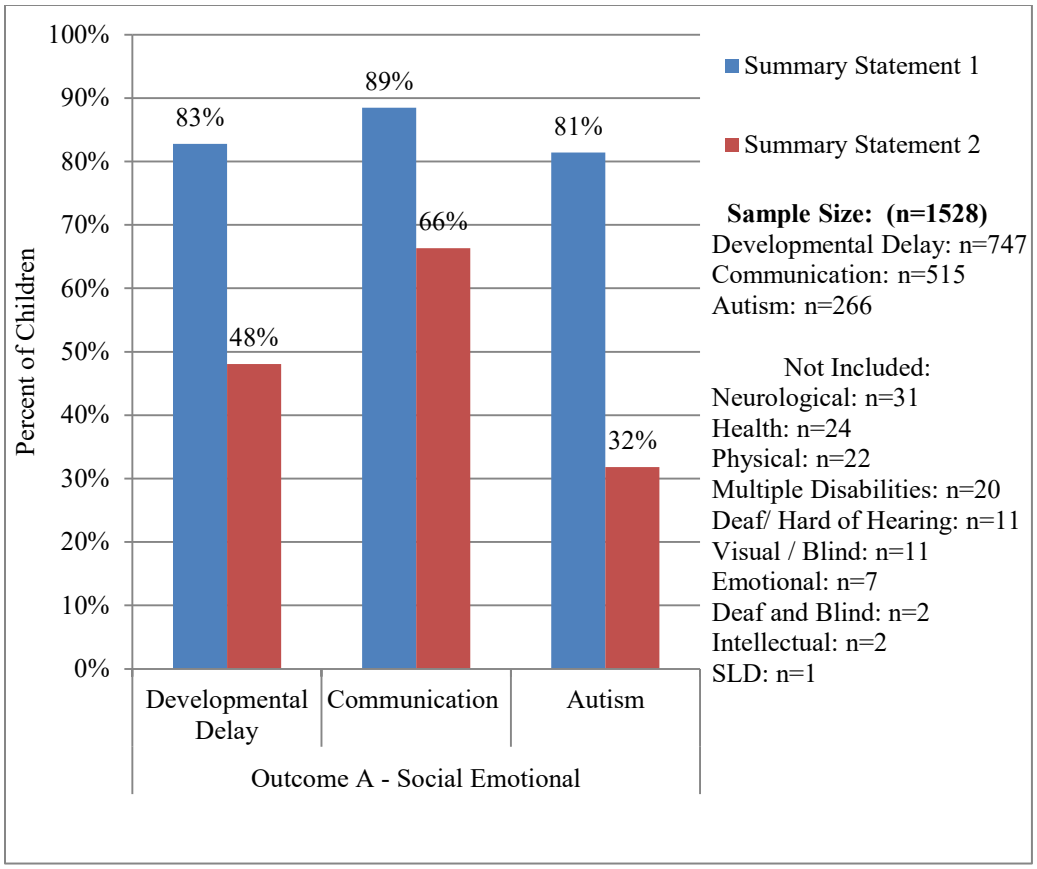


Figure 10. Summary Statement Values by Disability, 2012 & 2013 Exiting Children.

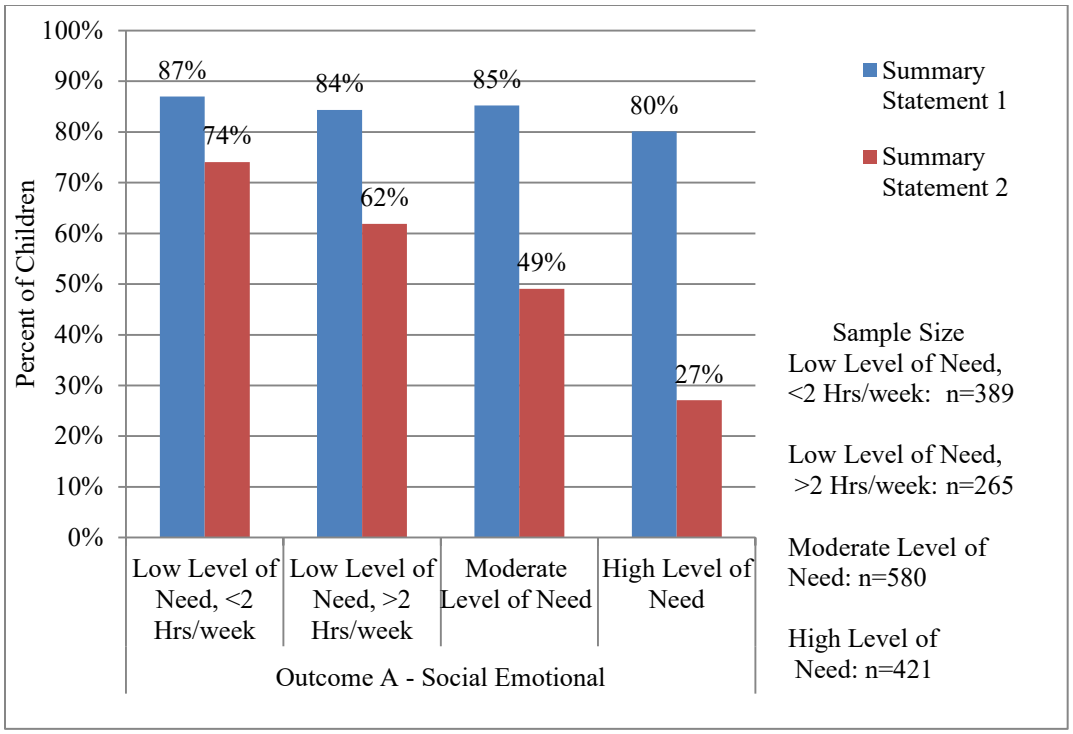


Figure 11. Summary Statement Values by Identified Level of Need, 2012 & 2013 Exiting Children.

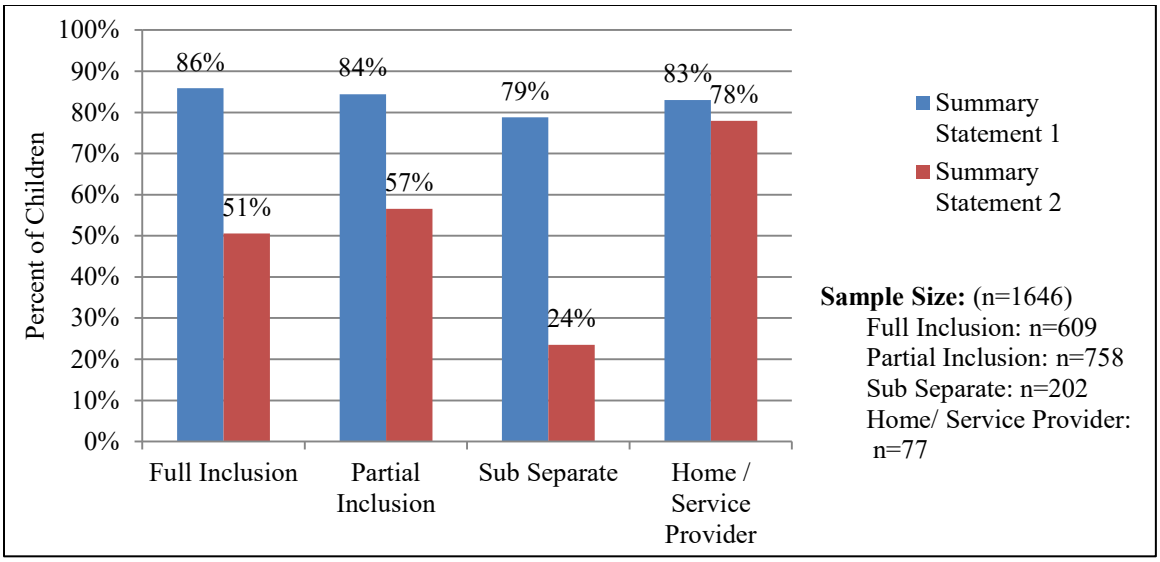


Figure 12. Summary Statement Values by Placement, 2012 & 2013 Exiting Children.

Table 1

Overview of the Selected Studies

Study	Design	Purpose	Participants	Nature of Intervention	Results
Benedict, Horner, & Squires, 2007	Multiple Baseline (2 Studies)	Assess implementation of PBS in a region	15 EC Settings (6 Head Start, 6 community preschools, 3 special education)	Assessment of implementation of features of PBS across classrooms using Pre-SET	Mean percentage of implementation across classrooms was 30.79% with a range of 13.33%-46.48% and a standard deviation of 11.80
		Evaluate the impact of PBS consultation on teacher's use of PBS & children's behavior	Four preschool classrooms (children ages 3-6)	4 classrooms selected to participate in consultation based on results of pre-SET, including approximately 7 in class consultation sessions	Increases in the mean percentage of features of PBS implemented from preassessment to post assessment

Study	Design	Purpose	Participants	Nature of Intervention	Results
Blair, Fox, & Lentini, 2010	Multiple Baseline	Evaluate the effects of implementation of assessment-based behavior support plans on children with challenging behaviors	Three children with disabilities ages 3-4	Implementation of assessment-based behavior support plans by lead and assistant teachers	Increases in the use of PBS and intervention strategies by teachers. Decreases in problem behavior and increases in engagement during circle time for children
Carter & Van Norman, 2010	Multiple Baseline	Investigate the effects of consultation on teacher's implementation of universal PBS practices and children's academic engagement	Four preschool classrooms (children ages 2-5)	Teachers received an initial 1 hour consultation session with an action plan on their use of PBS and then participated in at least one additional consultation session	Strong relationship between consultation & implementation of PBS skills. High levels of academic engagement following consultation

Study	Design	Purpose	Participants	Nature of Intervention	Results
Carter, Van Norman, & Tredwell, 2011	Multiple Baseline	Document the implementation process for program-wide PBS in one early childhood program	One early childhood classroom (children 6 weeks to 5 years)	Implementation of program-wide PBS in one program serving 250 children	The leadership team supported implementation but teacher buy-in increased when the program began to see changes in student behavior.
Duda, Dunlap, Fox, Lentini, & Clark, 2004	A-B-A-B Design	Examine the effects of PBS on children under 4 years old	Two three year old girls in a community preschool	Coaching for teachers during baseline measurement, implementation of behavior support strategies, withdrawal of supports, then reintroduction of the intervention until stability was achieved	Higher rates of engagement and lower rates of challenging behaviors for both children during intervention phases

Study	Design	Purpose	Participants	Nature of Intervention	Results
Fox, Hemmeter, Snyder, Binder, & Clarke, 2011	Single subject, multiple probe	Study the impact of training and coaching on implementation of Pyramid Model practices	Three early childhood classroom teachers	Teachers participated in a workshop and coaching sessions using the TPOT until they met implementation fidelity. Follow up observations were conducted after fidelity was met	Increases in use of the practices were seen with coaching but all teachers still exhibited "red flag" behaviors after coaching
Hall et al., 2007	Pre-post single subject pilot study	Study the impact of PBS on family quality-of-life outcomes	Fourteen families whose children attended an early childhood center	Families of children with challenging behaviors completed pre and post assessments and participated in 12 PBS workshop sessions over one year	Overall measures of stress decreased for families and children showed improvement in measures of social emotional development

Study	Design	Purpose	Participants	Nature of Intervention	Results
Muscott, Mann, & LeBrun, 2008	Case Study	Evaluate the implementation of PBS statewide	Twenty eight early childhood and K-12 programs	26 programs applied for and participated in an implementation of PBS as part of a state-supported cohort	A majority of schools were able to achieve implementation fidelity within 2 years and maintain that fidelity for a third year. Programs saw a reduction in office discipline referrals and suspensions. Implementation fidelity was also associated with gains in math scores.
Muscott, Pomerleau, Szczesiul, 2009	Case Study	Evaluate the implementation of PBS statewide	Sixteen early childhood programs at forty seven different sites	Description of the challenges facing sites as they participate in a statewide implementation of PBS	Most programs were successful in implementing preventative features of PBS but had difficulty with challenging behavior features.

Study	Design	Purpose	Participants	Nature of Intervention	Results
Stoiber & Gettinger, 2011	Multiple Baseline	Experimental analysis of teachers' use of functional assessment and PBS for addressing challenging behaviors	Thirty five teachers and ninety children aged 4-7 years	Teachers were coached in a five step procedure to design and implement functional assessments and PBS	Experimental teachers had increased resilience and use of PBS practices. Experimental children had fewer challenging behaviors and more positive behaviors.
Voorhees, Walker, Snell, & Smith, 2013	Multiple Baseline	Evaluate the use of individualized PBS	Three children in a Head Start Classroom	Teachers participated in classroom staff training, created functional behavioral assessments, and participated in initial and follow up coaching sessions	High implementation fidelity was measured in the classrooms and children showed decreases in inappropriate behaviors and increases in appropriate behaviors.

Table 2

Analysis of Experimental Studies of Positive Behavioral Supports in Early Childhood

	Study by Authors and Date										
Quality Indicator	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
Quality Indicators for Describing Participants											
1. Was sufficient information provided to determine/confirm whether the participants demonstrated the disability(ies) or difficulties presented?	0	1	0	1	1	1	1	1	1	1	1
2. Were appropriate procedures used to increase the likelihood that	1	1	1	0	0	0	0	0	0	1	1

	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
Quality Indicator relevant characteristics of participants in the sample were comparable across conditions? 3. Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?	0	1	1	0	1	1	1	1	1	1	1
Quality Indicators for Implementation of the Intervention and Description of											

Quality Indicator	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
Comparison Conditions											
1. Was the intervention clearly described and specified?	1	1	1	1	1	1	1	1	1	1	1
2. Was the fidelity of implementation described and assessed?	1	1	1	0	1	1	1	1	1	1	1
3. Was the nature of services provided in comparison conditions described?	0	0	1	0	0	0	0	0	0	1	0
Quality Indicators for Outcome Measures											

Quality Indicator	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
1. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalized performance?	1	1	1	0	1	1	1	1	0	1	1
2. Were outcomes for capturing the intervention's effect measured at the appropriate times?	1	1	1	1	1	1	1	1	0	1	1
Quality Indicators for Data Analysis											

Quality Indicator	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
1. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the unit of analysis in the study?	1	1	1	1	1	1	1	0	1	1	1
2. Did the research report include not only inferential statistics but also effect size calculations?	1	1	1	0	0	0	0	1	0	1	1
Desirable Quality Indicators											

Quality Indicator	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
1. Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?	1	1	1	1	1	1	1	1	1	1	1
2. Did the study provide not only internal consistency reliability but also test–retest reliability and interrater reliability (when appropriate)	1	1	1	0	1	1	0	0	0	1	0

Quality Indicator	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?											
3. Were outcomes for capturing the intervention's effect measured beyond an immediate posttest?	0	1	0	0	0	1	0	1	1	0	0
4. Was evidence of the criterion-related validity and construct validity of the	1	1	1	1	1	1	1	1	1	1	1

	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
Quality Indicator measures provided? 5. Did the research team assess not only surface features of fidelity implementation (e.g., number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?	1	1	1	1	1	1	1	1	1	1	1

Quality Indicator	(Benedict, Horner, & Squires, 2007)	(Blair, Fox, & Lentini, 2010)	Carter & Van Norman, 2010)	(Carter, Van Norman, & Tredwell, 2011)	(Duda, Dunlap, Fox, Lentini, & Clark, 2004)	(Fox, Hemmeter, Snyder, Binder, & Clarke, 2011)	(Hall et al., 2007)	(Muscott, Mann, & LeBrun, 2008)	Muscott, Pomerleau, Szczesiul, 2009)	(Stoiber & Gettinger, 2011)	(Voorhees, Walker, Snell, & Smith, 2013)
6. Was any documentation of the nature of instruction or series provided in comparison conditions?	0	0	0	0	0	0	0	0	0	1	0
7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?	0	0	0	0	0	0	0	0	0	0	0
8. Were results presented in a clear, coherent fashion?	1	1	1	1	1	1	1	1	1	1	1

Note. Adapted from "Quality Indicators for Group Experimental and Quasi-Experimenal Research in Special Education," by R. Gersten, L.S. Fuchs, D. Compton, M. Coyne, C. Greenwood, and M. S. Innocenti, 2005, *Exceptional Children, 71,* p. 152.

Table 3

Data Collection Tools and Indicators of High, Moderate, and Low Fidelity

Data Collection Tool	Description of Tool	Data Source	Frequency	Individual Responsible	High Fidelity	Moderate Fidelity	Low Fidelity
State Systemic Improvement Plan (SSIP) Phase I & Phase II Reports	Federal report submitted annually to OSEP that articulates the plan for implementation of the Pyramid Model in MA	State Performance Plan / Annual Performance Report (SPP/APR)	Annual	Principal Investigator	Full Alignment with OSEP requirements, all components present	Alignment with OSEP requirements, some components missing	Misalignment with OSEP requirements and/or several components (more than 4) missing
Training Materials	The PowerPoints, handouts, and presentations provided to participants	Training materials and training evaluations	As trainings occur	Principal Investigator	Training materials are existing Pyramid Model resources, participants indicate high-quality presentations	Most training materials are existing Pyramid Model resources, participants indicate quality presentations but unanswered questions	Training uses few if any Pyramid Model materials, evaluations indicate confusion or lack of understanding

Data Collection Tool	Description of Tool	Data Source	Frequency	Individual Responsible	High Fidelity	Moderate Fidelity	Low Fidelity
Interviews	Interviews and ongoing discussions with National Pyramid Model Consortium experts	National Pyramid Model Consortium Consultants	As needed, at least monthly	Principal Investigator	National consultants indicate the model is being implemented as designed and with fidelity	National consultants indicate the model is being implemented as designed with some minor modifications, possibility of lack of full implementation fidelity	National consultants indicate the model is being implemented with significant modifications that may result in a loss of fidelity
Training Participant Data	Information collected as part of the training registration process including participant's role, district, and participation in previous training activities	MA ESE Training Registration Lists	As trainings occur	Principal Investigator	At each training, the intended audience is present	The intended audience is partly present at each training with some individuals missing or in attendance with inappropriate roles	The intended audience is mostly absent or the audience consists of individuals without the proper role or background

Data Collection Tool	Description of Tool	Data Source	Frequency	Individual Responsible	High Fidelity	Moderate Fidelity	Low Fidelity
District/Program Benchmarks of Quality	A self-assessment on which district indicate which critical features of the Pyramid Model are in place, partially in place, or not in place	District Leadership Teams and External Coaches	At least 3 times per school year	District Leadership Teams, External Coaches, and MA ESE Staff	At least 75 percent of the critical features are "in place" as reported on the tool	Between 40 and 74 percent of the critical features are "in place" as reported on the tool	Less than 40 percent of the critical features are "in place" as reported on the tool
Teaching Pyramid Observation Tool (TPOT)	A published measure of implementation fidelity in the classroom	Classroom-based Coaches and External Coaches	At least twice a year	Classroom-based Coaches, External Coaches, and MA ESE Staff	A score of 80% or higher with no red flags indicated on the tool	A score of 50-80% with less than 4 red flags indicated on the tool	A score of less than 50% and/or 4 or more red flags indicated
External Coach Contact Record	Online tool for External Coaches to report activities, trainings, observations, and	External Coaches	Each time there is a substantive contact with district personnel	External Coaches and MA ESE Staff	External Coaches provide qualitative data indicating high fidelity of	External Coaches provide qualitative data indicating moderate fidelity of	External Coaches provide qualitative data indicating low fidelity of implementati

Data Collection Tool	Description of Tool	Data Source	Frequency	Individual Responsible	High Fidelity	Moderate Fidelity	Low Fidelity
	consultations that occur at the district and classroom level				implementation based on criteria established by the Pyramid Model Consortium without any concerns	implementation based on criteria established by the Pyramid Model Consortium with limited concerns	on based on criteria established by the Pyramid Model Consortium with significant concerns

Table 4
Measures of Fidelity at the State, District, and Classroom Level by Dimension of Fidelity

Dimension of Fidelity	Criteria Definition	Measure of Fidelity at the State Level	Measure of Fidelity at the District/ Program Level	Measure of Fidelity at the Classroom Level
Adherence	Are the components of the intervention being delivered as designed?	SSIP Phase I & Phase II Reports Training Materials Interviews	Training Materials Training Participant Data District Benchmarks of Quality External Coach Contact Record	TPOT External Coach Contact Record
Duration	How many sessions are being delivered? At what frequency? At what length?	SSIP Phase I & Phase II Reports Training Materials	Training Materials External Coach Contact Record	TPOT

Dimension of Fidelity	Criteria Definition	Measure of Fidelity at the State Level	Measure of Fidelity at the District/ Program Level	Measure of Fidelity at the Classroom Level
Quality of Delivery	Are the intended methods and resources being used to deliver the intervention consistent with how the intervention was designed?	Training Materials Training Participant Data Interviews	District Benchmarks of Quality External Coach Contact Record	TPOT External Coach Contact Record
Participant Responsiveness	How engaged in and involved with the implementation of the intervention are participants?	Training Materials Training Participant Data	District Benchmarks of Quality External Coach Contact Record	External Coach Contact Record
Program Differentiation	Are the critical features that distinguish this program from the control condition in place during implementation?		District Benchmarks of Quality	TPOT

Note. SSIP = State Systemic Improvement Plan; TPOT = Teaching Pyramid Observation Tool. Dimensions of fidelity as defined by Dusenbury et al., (2003) and O'Donnell (2008).

Table 5

Tools and the Elements of Each Tool Informing by Dimension of Fidelity

Dimension of Fidelity	% of Tools Rated as High Fidelity	% of Tools Rated as Moderate Fidelity	% of Tools Rated as Low Fidelity	Tool	Tool Element	Fidelity Rating
Adherence	33.3%	50%	16.7%	SSIP Report	Alignment with implementation plan and OSEP requirements	High
				Interviews	Feedback on adherence to model as developed	Moderate
				Training Data	Training materials and schedule of trainings	High
				District Benchmarks of Quality	Results of Benchmarks of Quality Analysis	Moderate
				TPOT	Scores from recent TPOTs	Low
				External Coach Contact Record	Description of implementation	Moderate
Dose/ Duration	60%	0%	40%	SSIP Report	Alignment with implementation plan	High
				Interviews	Feedback on frequency and length of training sessions	High

Dimension of Fidelity	% of Tools Rated as High Fidelity	% of Tools Rated as Moderate Fidelity	% of Tools Rated as Low Fidelity	Tool	Tool Element	Fidelity Rating
Quality of Delivery	80%	20%	0%	Training Data	The type, number, and frequency of trainings	High
				TPOT	The frequency with which TPOTs are administered	Low
				External Coach Contact Record	The frequency with which coaching occurs	Low
				SSIP Report	Description of methods and resources used to deliver the intervention	Moderate
				Interviews	Feedback on methods and resources used to deliver the intervention	High
				Training Data	Training materials	High
				TPOT	TPOT reliability training data	High
				External Coach Contact Record	Methods and resources being used to implement the Pyramid Model	High
Participant Responsiveness	40%	40%	20%	SSIP Report	Participant engagement data	High
				Training Data	Representativeness and role of participants from implementing	High

Dimension of Fidelity	% of Tools Rated as High Fidelity	% of Tools Rated as Moderate Fidelity	% of Tools Rated as Low Fidelity	Tool	Tool Element	Fidelity Rating
					districts	
				Benchmarks of Quality	Frequency of administration	Moderate
				TPOT	Frequency of administration	Low
				External Coach Contact Record	Engagement data	Moderate
Program Differentiation	N/A	N/A	N/A	N/A	N/A	N/A

Table 6

Timeline for Pyramid Model Implementation & Data Collection Activities: 2015-2016 School Year

Date	Activity/Training	Audience	Data Collection Tool Utilized
Ongoing	Interviews with Pyramid Model Consortium Staff	N/A	Interview
Ongoing	External Coach Contact Records completed as substantive contacts with districts occur	N/A	External Coach Contact Record
4/1/15	SSIP Phase I Report Submitted to OSEP	OSEP	SSIP Reports
5/20-5/22/15	Leadership Team Academy (2.5 Days) (5 Districts participated)	District Leadership Team and External Coaches	Training Data District Benchmarks of Quality
6/29-6/30/15	Pyramid Practices Training Part 1 (2 Days) (Approximately 30 practitioners participated)	Practitioners (teachers), Classroom Coaches, Behavior Specialists, and External Coaches	Training Data
7/1/15	SSIP Phase I Report Feedback from OSEP	OSEP, MA ESE	SSIP Reports
8/18-8/19/15	Pyramid Practices Training Part 2 (2 Days) (Approximately 30 practitioners participated)	Practitioners (teachers), Classroom Coaches, Behavior Specialists, and External Coaches	Training Data
9/22/15	Kick Off Webinar – District leadership teams from May 2015	District leadership teams and External Coaches	Training Data
9/30-10/2/15	Leadership Team Academy (2.5 Days) 4 Districts participated	District leadership team and External Coaches	Training Data District Benchmarks of Quality
10/2/15	Positive Solutions Train-the-Trainer 17 Individuals participated	External Coaches and Community Trainers	Training Data
10/21-10/22/15	TPOT (Teaching Pyramid Observation Tool) Training (2 Days) 28 participants including	Classroom Coaches and External Coaches	Training Data

Date	Activity/Training	Audience	Data Collection Tool Utilized
	community partners		
10/27-10/29/15	Leadership Team Academy (2.5 Days) 10 Districts participated	District leadership team and External Coaches	Training Data District Benchmarks of Quality
11/5, 11/12, 11/19, 12/3, & 12/10/15	Five, 1-hour coaches training calls – Virtual Approximately 10 people participating	Classroom Coaches and External Coaches	Training Data
11/12-11/13/15	Pyramid Practices Training Part 1 (2 Days) Approximately 90 participants	Practitioners (teachers), Classroom Coaches, Behavior Specialists, and External Coaches	Training Data
1/14-1/15/16	Pyramid Practices Training Part 2 (2 Days) Approximately 90 participants	Practitioners (teachers), Classroom Coaches, Behavior Specialists, and External Coaches	Training Data
2/1-2/2/16	TPOT (Teaching Pyramid Observation Tool) Training (2 Days) Approximately 30 classroom coaches trained	Classroom Coaches and External Coaches	Training Data
2/26/2016	Prevent/Teach/Reinforce Young Children (PTR-YC) Training (1 Day) Approximately 30 participants	Behavior Specialists and External Coaches	Training Data
3/1/16	Leadership Team Mid-Year Meeting (1 Day) 16 Districts Participated	District leadership teams and External Coaches	Training Data District Benchmarks of Quality
3/3-4/14/16	Five, 1-hour coaches training calls – Virtual Approximately 15 individuals participated	Classroom Coaches and External Coaches	Training Data
3/14-3/15/16	Pyramid Practices Training Part 1 (2 Days)	Practitioners (teachers), Classroom Coaches, Behavior	Training Data

Date	Activity/Training	Audience	Data Collection Tool Utilized
		Specialists, and External Coaches	
4/1/16	SSIP Phase II Report	OSEP	SSIP Reports
4/25/16	Pyramid Practices Training Part 2 (1 Day) Approximately 30 participants	Practitioners (teachers), Classroom Coaches, Behavior Specialists, and External Coaches	Training Data
4/26/16	Prevent/Teach/Reinforce Young Children (PTR-YC) Training (1 Day) Approximately 30 participants	Behavior Specialists and External Coaches	Training Data
6/2/16	Leadership Team End of Year Meeting (1 Day) 16 Districts Participated	District Leadership Teams and External Coaches	Training Data District Benchmarks of Quality
7/1/16	SSIP Phase II Report Feedback from OSEP	MA ESE, OSEP	SSIP Reports
9/2016	Collection of 2015-2016 TPOT Data from participating districts	TPOT	TPOT
9/26/16	Beginning of the Year Leadership Team Meeting	District Leadership Teams and External Coaches	Training Data

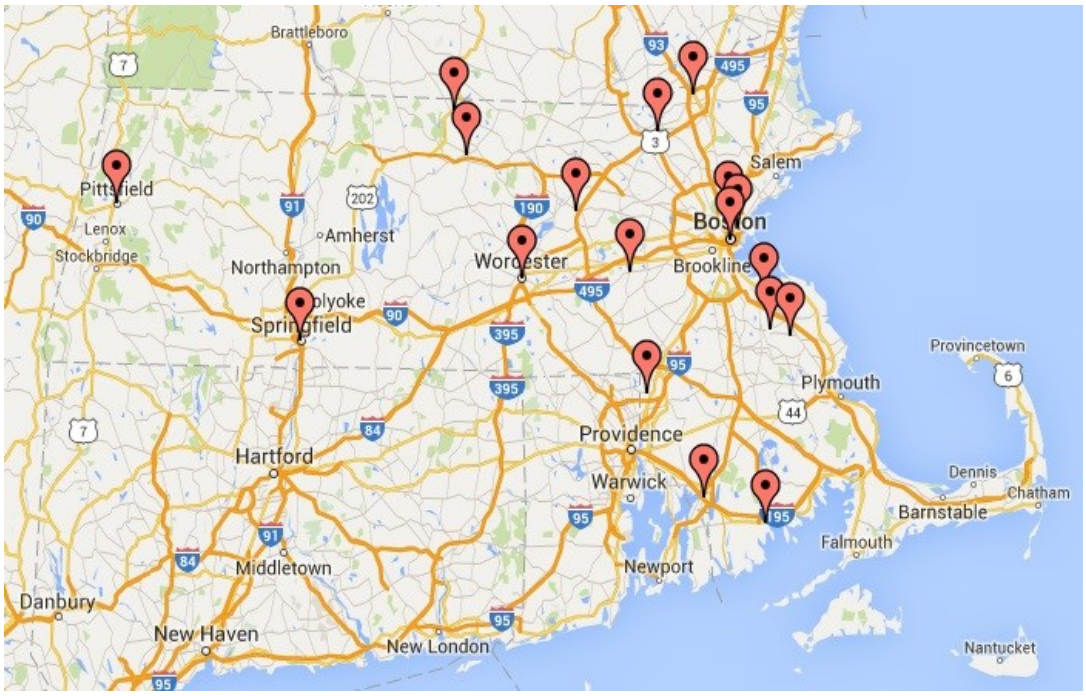


Figure 13. Map of Participating Pyramid Model Initiative Districts in Massachusetts

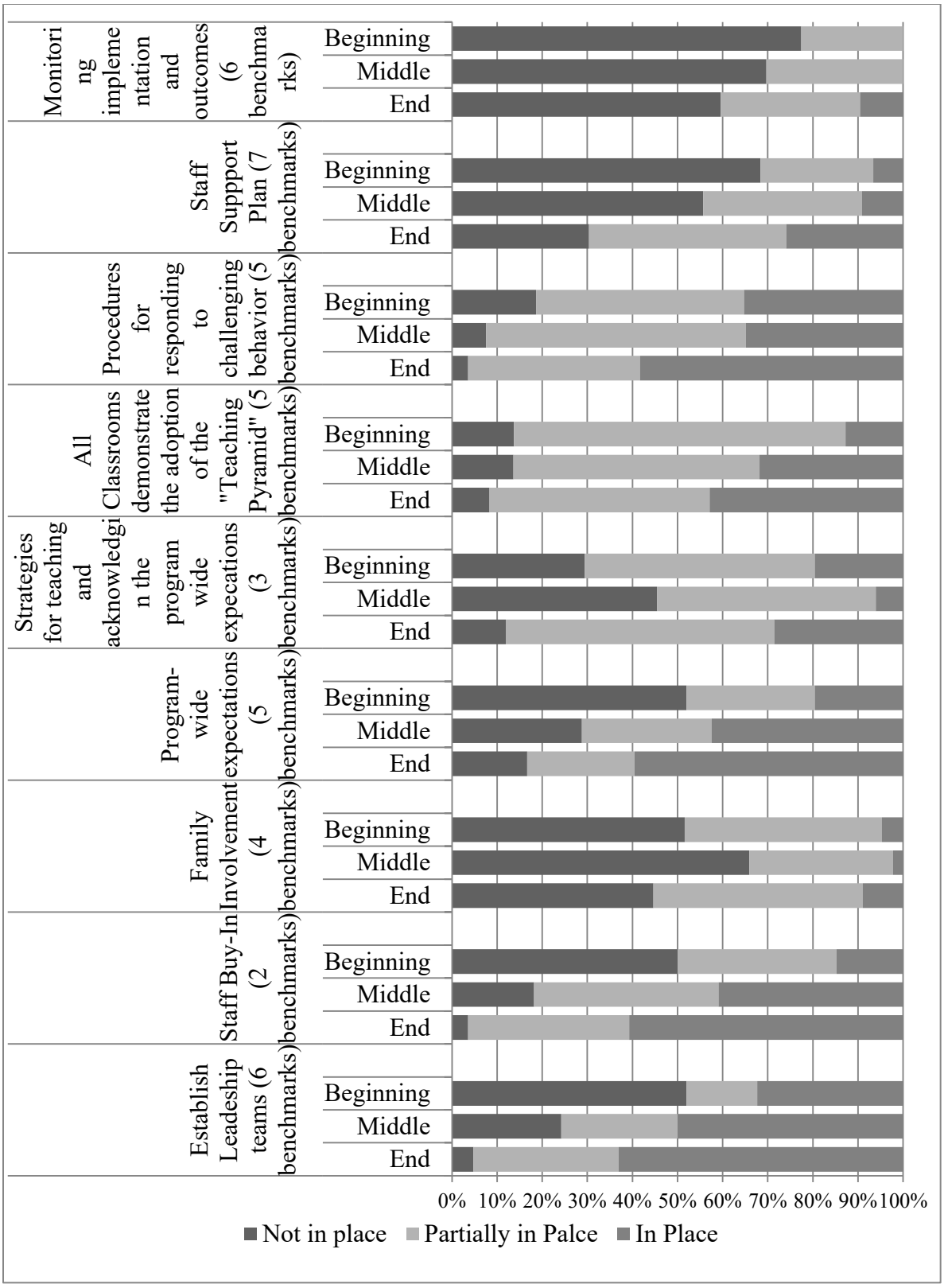


Figure 14. Aggregate Benchmarks of Quality for Participating Districts: Beginning, Middle, and End of Year Administration.

Table 7
Results Matrix: Level of Fidelity as Measured by Each Data Collection Tools on Relevant Dimensions of Fidelity

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
Adherence (Are the components of the intervention being delivered as designed?)	High (3 points)	*Full alignment with Phase I Implementation Plan for EBP and OSEP Requirements*	Interviews with developers of model indicate full adherence without any concerns	*Training materials are consistent with published materials for the EBP and are being delivered in the correct sequence to the correct audience*	Participating districts indicate at least 80% of benchmarks are "in place" on most recently completed Benchmarks of Quality	The average score for all TPOTs completed in the last 3 months across participating districts is 80% or higher	External coaches report that the components of the intervention are being delivered as designed with minimal variance and no concerns about adherence to the model in at least 80% of participating districts (15 out of 18)
	Moderate (2 Points)	Alignment with Phase I Implementation Plan and OSEP Requirements, some components are missing	*Interviews with developers of model indicate adherence with some minor variations or concerns*	Training materials are consistent with published materials for the EBP with some modifications and/or omissions and are generally being delivered in the correct sequence to the correct audience with some	*Participating districts indicate between 40 and 79% of benchmarks are "in place" on most recently completed Benchmarks of Quality*	The average score for all TPOTs completed in the last 3 months across participating districts is between 50 and 79%	*External coaches report that the components of the intervention are mostly being delivered as designed but there are some concerns about adherence to the model in no more than 50% of participating districts.*

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
				exceptions			
	Low (1 Points)	Misalignment with Phase I implemetnation Plan, OSEP Requirements, and/or more than 4 components missing for EBP	Interviews with developers of model indicate lack of adherence with significant variations or concerns about implementation	Training materials are inconsistent with published materials for the EBP with modifications and/or omissions and are generally not being delivered in the correct sequence to the correct audience	Participating districts indicate less than 40% of benchmarks are "in place" on most recently completed Benchmarks of Quality	*The average score for all TPOTs completed in the last 3 months across participating districts is less than 50%*	External coaches report that some compoents of the intervention are being delivered as designed but there are concerns about adherence to the model in more than half of the participating districts

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
Duration or Dose (How many sessions are being delivered? At what frequency? At what length?)	High (3 points)	*Full alignment with Phase I Implementation Plan for EBP*	*Interviews with developers of model indicate sessions are being delivered at the appropriate frequency and length.*	*Data on the type, number, length, and frequency of trainings indicates alignment with the EBP*		TPOTs are completed by district internal coaches for all teachers implementing this model at least twice a year for 80% or more of participating districts	External coaches are providing on-site and/or virtual coaching to district leadership teams at least every other month in more than 80% of participating districts
	Moderate (2 Points)	Alignment with Phase I Implementation Plan, some components are missing for EBP	Interviews with developers of model indicate sessions are being delivered at the appropriate frequency and length with some minor	Data on the type, number, length, and frequency of trainings indicates alignment with the EBP with minor substantive variations or concerns		TPOTs are completed by district internal coaches for all teachers implementing this model at least twice a year for less than 80% but more than 50% of participating districts	External coaches are providing on-site and/or virtual coaching to district leadership teams at least every other month for less than 80% but more than 50% of participating districts

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
			variations or concerns				
	Low (1 Point)	Misalignment with Phase I implemetnation Plan and/or more than 4 components missing for EBP	Interviews with developers of model indicate sessions are being delivered at the appropriate frequency and length with significant variations or concerns about implement	Data on the type, number, length, and frequency of trainings indicates alignment with the EBP with significant variations or concerns		*TPOTs are completed by district internal coaches for all teachers implementing this model at least twice a year for less than 50% of participating districts*	*External coaches are providing on-site and/or virtual coaching to district leadership teams at least every other month in less than 50% of participating districts*

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews ation	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
Quality of Delivery (Are the intended methods and resources being used to deliver the intervention consistent with how the intervention was designed?)	High (3 points)	The description of the methods and resources used to deliver the intervention indicates full alignment with the EBP as designed without modifications	*Interviews with developers of model indicate the methods and resources being used are consistent with how the intervention was designed*	*A review of training materials indicates that they are fully aligned with the with the resources developed by the intervention's creators and are being used in a manner consistent with their design*		*Individuals administering the TPOT have been assessed and found reliable on the use of this tool*	*External coaches report that the methods and resources being used to deliver the intervention are consistent with how the intervention was designed in 80% or more of participating districts*

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
	Moderate (2 Points)	*The description of the methods and resources used to deliver the intervention indicates alignmnent with the EBP as designed with minor modifications*	Interviews with developers of model indicate the methods and resources being used are consistent with how the interventio n was designed with some minor variations or concerns	A review of training materials indicates that they are fully aligned with the with the resources developed by the intervention's creators and are being used in a manner consistent with their design with some minor substantive variations or concerns		Most but not all individuals administering the TPOT have been assessed and found reliable on the use of this tool	External coaches report that the methods and resources being used to deliver the intervention are consistent with how the intervention was designed in less than 80% and more than 50% of participating districts

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
	Low (1 Points)	The description of the methods and resources used to deliver the intervention indicates lack of alignment with the EBP as designed with significant modifications	Interviews with developers of model indicate the methods and resources being used are inconsistent with how the intervention was designed	A review of training materials indicates that they are fully aligned with the with the resources developed by the intervention's creators and are being used with significant substantive variations or concerns about their quality		Few, if any, individuals administering the TPOT have been assessed and found reliable on the use of this tool	External coaches report that the methods and resources being used to deliver the intervention are consistent with how the intervention was designed in less than 50% of participating districts
Participant Responsiveness (How engaged in and involved with the implementation of the intervention are participants	High (3 points)	*The SSIP Reports indicate that participating districts are fully engaged with the implementation of the intervention*		*Training activities have representation from at least 80% of participating districts on average and the appropriate individuals attend each training*	The Benchmarks of Quality are completed by district leadership teams at least twice a year for 80% or more of	TPOTs are completed by district internal coaches for all teachers implementing this model at least twice a year for 80% or more of participating districts	External coaches indicate that at least 80% of districts are actively engaged in implementation activities in both classrooms and across the program/district through the use of Pyramid Model practices, the

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
?)					participating districts		development of new resources, regular leadership team meetings, and/or family engagement around the model
	Moderate (2 Points)	The SSIP Reports indicate that participating districts are somewhat engaged with the implementation of the intervention		Training activities have representation from at least 50% of participating districts on average and/or there are some attendees at trainings who do not have the appropriate background or prerequisite knowledge to participate	*The Benchmarks of Quality are completed by district leadership teams at least twice a year for less than 80% but more than 50% of participating districts*	TPOTs are completed by district internal coaches for all teachers implementing this model at least twice a year for less than 80% but more than 50% of participating districts	*External coaches indicate that at least 50% but less than 80% of districts are actively engaged in implementation activities in both classrooms and across the program/district through the use of Pyramid Model practices, the development of new resources, regular leadership team meetings, and/or*

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
							family engagement around the model

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
	Low (1 Point)	The SSIP Reports indicate that participating districts are fully engaged with the implementation of the intervention		Training activities have representation from at less than 50% of participating districts on average and/or a majority of attendees at trainings do not have the appropriate background or prerequisite knowledge to participate	The Benchmarks of Quality are completed by district leadership teams at least twice a year by less than 50% of participating districts	*TPOTs are completed by district internal coaches for all teachers implementing this model at least twice a year forless than 50% of participating districts*	External coaches indicate that less than 50% of participating districts are actively engaged in implementation activities in both classrooms and across the program/district through the use of Pyramid Model practices, the development of new resources, regular leadership team meetings, and/or family engagement around the model
<u>Program Differentiation (Are the critical features that distinguish this program</u>	High (3 points) Moderate (2 Points)						

Dimension of Fidelity	Level of Fidelity	SSIP Reports	Interviews	Training Data	District Benchmarks of Quality	TPOT	External Coach Contact Record
from the control condition in place during implementation?)	Low (1 Point)						

Note. Identified responses for this study are presented in italics. If a portion of the table is blank then that tool does not inform that dimension of fidelity.

Table 8

Results Matrix: Number and Percentage of Tools at High, Moderate and Low Fidelity

Data Collection Tool	Implementation Fidelity		
	High	Moderate	Low
	N (%)	*N (%)*	*N (%)*
SSIP Reports	3 (75%)	1 (25%)	0 (0%)
Interviews	2 (67%)	1 (33%)	0 (0%)
Training Data	4 (100%)	0 (0%)	0 (0%)
District Benchmarks of Quality	0 (0%)	2 (100%)	0 (0%)
TPOT	1 (25%)	0 (0%)	3 (75%)
External Coach Contact Record	1 (25%)	2 (50%)	1 (25%)

Table 9

Results Matrix: Number and Percentage of Tools at High, Moderate and Low Fidelity for Each Relevant Dimensions of Fidelity

Dimension of Fidelity	Implementation Fidelity			
	High	Moderate	Low	Summary
	N (%)	*N (%)*	*N (%)*	*N (%)*
Adherence	2 (33.3%)	3 (50%)	1 (16.7%)	13/18 (72.2%)
Dose/Duration	3 (60%)	0 (0%)	2 (40%)	11/15 (61.1%)
Quality	4 (80%)	1 (20%)	0 (0%)	14/15 (93.3%)
Participant Responsiveness	2 (40%)	2 (40%)	1 (20%)	11/15 (73.3%)
Program Differentiation	N/A	N/A	N/A	
Summary				49/63 (77.8%)

Table 10

Statewide Trainings May 2015-September 2016

Training Name	Times Offered	Duration	Total Number of Attendees	Number of Districts Represented (% of Districts Represented)
Initial Leadership Team Meeting	3	2.5 Days (17.5 Hours)	92	18 (100%)
Midyear Leadership Team Meeting	1	1 Day (7 Hours)	67	17 (94%)
End-of-Year Leadership Team Meeting	1	1 Day (7 Hours)	52	14 (78%)
Pyramid Model Practices Training (Modules 1, 2, 3A, & 3B)	3	4 Days (28 Hours)	172	18 (100%)
Teaching Pyramid Observation Tool (TPOT)	2	2 Days (14 Hours)	59	15 (83%)
Positive Solutions Train-the-Trainer for External Coaches	1	1 Day (7 Hours)	16	N/A
Prevent, Teach, Reinforce, Young Children (PTR-YC)	1	1 Day (7 Hours)	43	16 (89%)
Beginning of the Year Leadership Team Meeting	1	1 Day (7 Hours)		14 (78%)

External Coaches Pyramid Model Practices Coaching Calls	6	1 Hour Each	8	N/A
External Coaches Pyramid Model Practices Co-Training with the National Trainer	8	Half Day per Practices Training (Approximately 8 hours for each coach)	8	N/A

.

Appendix A: Acronyms and Abbreviations

COS Process	The Child Outcomes Summary Process
CSEFEL	Center for the Social Emotional Foundations of Early Learning
CWD	Children with Disabilities
ECTA	The Early Childhood Technical Assistance Center
ECO	The Early Childhood Outcomes Center
DCF	Department of Children and Families
DMH	Department of Mental Health
DPH	Department of Public Health
ECSE	Early Childhood Special Education
EEC	Department of Early Education and Care
EOE	Executive Office of Education
Indicator 17	The State Systemic Improvement Plan
Indicator 7	Early Childhood Outcomes Measurement
LEA	Local Education Agency
MA ESE	Massachusetts Department of Elementary and Secondary Education
OSEP	Federal Office of Special Education Programs
PBS	Positive Behavior Supports
PTR-YC	Prevent, Teach, Reinforce, Young Children
SEPP	Special Education Policy and Planning
SSIP	State Systemic Improvement Plan
SWD	Students with Disabilities
TPOT	Teaching Pyramid Observation Tool

Appendix B: Data Collection Tools
Early Childhood Program-Wide PBS Benchmarks of Quality

Program Name: _____ Location: _____
Date: _____

Team Members:

_____
_____
_____

For each item note whether the critical element is "Not in Place" (0 points), "Partially in Place" (1 point), or "In Place" (2 points).

Critical Element: Establish leadership team
Benchmarks of Quality:

1. Team has broad representation that includes at a minimum a teacher, administrator and a member with expertise in behavior support. Other team members might include parent, teaching assistant, related service specialists and other program personnel.

2. Team has administrative support. Administrator attends meetings and trainings, is active in problem-solving to ensure the success of the initiative, and is visibly supportive of the adoption of the model.

3. Team has regular meetings. Team meetings are scheduled at least 1x per month for a minimum of 1 hour. Team member attendance is consistent.

4. Team has established a clear mission/purpose. The team purpose or mission statement is written. Team members are able to clearly communicate the purpose of the leadership team.

5. Team develops an implementation plan that includes all critical elements. A written implementation plan guides the work of the team. The team reviews the plan and updates their progress at each meeting. Action steps are identified to ensure achievement of the goals.

6. Team reviews and revises the plan at least annually.

Critical Element: Staff Buy-In
Benchmarks of Quality:

7. Staff are aware of and supportive of the need for a program wide system for addressing children's social emotional development and challenging behavior. A staff poll establishes buy-in before the initiative is launched.

8. Staff input and feedback is obtained throughout the process - coffee break with the director, focus group, suggestion box. Leadership team provides update on the process and data on the outcomes to program staff on a regular basis.

9. Family input is solicited as part of the planning process. Families are Family Involvement informed of the initiative and asked to provide feedback on program-wide adoption and mechanisms for promoting family involvement in the initiative.

10. There are multiple mechanisms for sharing the program wide plan with families including narrative documents, conferences, and parent meetings to ensure that all families are informed of the initiative.

11. Family involvement in the initiative is supported through a variety of mechanisms including home teaching suggestions, information on supporting social development, and the outcomes of the initiative. Information is shared through a variety of formats (e.g., meetings, home visit discussions, newsletters, open house, websites, family friendly handouts, workshops, rollout events).

12. Families are involved in planning for individual children in a meaningful and proactive way. Families are encouraged to team with program staff in the development of individualized plans of support for children including the development of strategies that may be used in the home and community.

Critical Element: Program-wide expectations
Benchmarks of Quality:

13. 2-5 positively stated program wide expectations are developed.

14. Expectations are written in a way that applies to both children and staff. When expectations are discussed, the application of expectations to program staff and children is acknowledged.

15. Expectations are developmentally appropriate and linked to concrete rules for behavior within activities and settings.

16. All program staff are involved in the development of the expectations.

17. Expectations are shared with families and staff assist families in the translation of the expectations to rules in the home.

18. Expectations are posted in classrooms and in common areas in ways that are meaningful to children, staff and families.

Critical Element: Strategies for teaching and Acknowledging the Program Wide Expectations
Benchmarks of Quality:

19. Instruction on expectations is embedded within large group activities, small group activities, and within individual interactions with children.

20. A variety of teaching strategies are used: teaching the concept, talking about examples and non-examples, scaffolding children's use of the expectations in the context of ongoing activities and routines. Instruction on expectations and rules occurs on a daily basis.

21. Strategies for acknowledging children's use of the expectations are developmentally appropriate and used by all program staff including administrative and support staff (e.g., clerical, bus drivers, kitchen staff).

Critical Element: All Classrooms Demonstrate Implementation of the Pyramid Model
Benchmarks of Quality:

22. Teachers and program staff have strategies in place to promote positive relationships with children, each other, and families and use those strategies Pyramid Model.

23. Teachers and program staff have arranged environments, materials, and curriculum in a manner that promotes social-emotional development and guides appropriate behavior.

24. Teachers and program staff are proficient at teaching social and emotional skills within daily activities in a manner that is meaningful to children and promotes skill acquisition.

25. Teachers and program staff respond to children's problem behavior appropriately using evidence-based approaches that are positive and provide the child with guidance about the desired appropriate behavior.

26. Teachers and program staff provide targeted social emotional teaching to individual children or small groups of children who are at-risk for challenging behavior.

27. Teachers and program staff initiate the development of an individualized plan of behavior support for children with persistent challenging behavior.

Critical Element: Procedures for Responding to Challenging Behavior
Benchmarks of Quality:

28. Strategies for responding to problem behavior in the classroom are developed. Teachers use evidence-based approaches to respond to problem behavior in a manner that is developmentally appropriate and teaches the child the expected behavior.

29. A process for responding to crisis situations related to problem behavior is developed. Teachers can identify how to request assistance when needed. A plan for addressing the child's individual behavior support needs is initiated following requests for crisis assistance.

30. A process for problem solving with other teachers around problem behavior is developed. Teachers can identify a process that may be used to gain support in developing ideas for addressing problem behavior within the classroom (e.g., peer-support, classroom mentor meeting, brainstorming session).

31. A team-based process for addressing individual children with persistent challenging behavior is developed. Teachers can identify the steps for initiating the team-based process including fostering the participation of the family in the process.

32. An individual or individuals with behavioral expertise are identified for coaching staff and families throughout the process of developing and implementing individualized intensive interventions for children in need of behavior support plans.

33. Strategies for partnering with families when there are problem behavior concerns are identified. Teachers have strategies for initiating parent contact and partnering with the family to develop strategies to promote appropriate behavior.

Critical Element: Professional Development and Staff Support Plan
Benchmarks of Quality:

34. A plan for providing ongoing support, training, and coaching in each classroom on the Pyramid Model practices is developed and implemented.

35. A data-driven coaching model is used to assist classroom staff with implementing the Pyramid Model practices to fidelity.

36. Staff responsible for facilitating behavior support processes are identified and trained.

37. A needs assessment is conducted with staff to determine training needs on the adoption of the Pyramid Model.

38. Individualized professional development plans are developed with all staff.

39. Group and individualized training strategies are identified and implemented.

40. Plans for training new staff are identified and developed.

41. Incentives and strategies for acknowledging staff are identified.

Critical Element: Monitoring Implementation and Outcomes
Benchmarks of Quality:

42. Process for measuring implementation fidelity is used.

43. Process for measuring outcomes is developed.

44. Data are collected and summarized.

45. Data are shared with program staff and families.

46. Data are used for ongoing monitoring, problem solving, ensuring child response to intervention, and program improvement.

47. Implementation Plan is updated/revised as needed based on the ongoing data.

Preschool PBS through Pyramid Strategies External Coach Contact Record

---

Pyramid Model Coach Contact Record

1) External Coach (CSPD Trainer):*

First Name:

Last Name:

Email Address:

2) District/Program Information*

District/Program Name:

School Name (if applicable):

Validation: %s format expected

3) What was the date of contact?*

4) Contact with:*

- ☐ Leadership Team
- ☐ Program Administrator
- ☐ Classroom Coach
- ☐ Behavior Specialist
- ☐ Data Coordinator
- ☐ SEPAC
- ☐ Family Organization
- ☐ Other - Write In (Required): *

5) Contact Method:*

- ☐ Site Visit (program, classroom, or home)
- ☐ Phone Call
- ☐ Email
- ☐ Off-site meeting
- ☐ Training
- ☐ Other - Write In (Required): _____ *

Logic: Show/hide trigger exists.
6) What types of support were provided? (check all that apply)*

- ☐ Leadership Team Support
- ☐ Classroom Coach Support
- ☐ Data Decision Making
- ☐ Behavior Systems
- ☐ Family Group Training

Logic: Hidden unless: Question "What types of support were provided? (check all that apply)" #6 is one of the following answers ("Leadership Team Support")
7) What types of Leadership Team support were provided?*

- ☐ Assist team leader in building meeting agenda/processes
- ☐ Attended meeting(s)
- ☐ Product development and review
- ☐ Assist with locating resources
- ☐ Assist with professional development / coaching
- ☐ Other - Write In (Required): _____ *

Logic: Hidden unless: Question "What types of support were provided? (check all that apply)" #6 is one of the following answers ("Classroom Coach Support")
8) What type(s) of classroom coach support were provided?*

- ☐ Observe teacher with internal coach
- ☐ Debrief with classroom coach
- ☐ Assist with coaches' professional development

- ☐ TPOT
- ☐ Other - Write In (Required): _____ *

Logic: Hidden unless: Question "What types of support were provided? (check all that apply)" #6 is one of the following answers ("Data Decision Making")
9) What type of data decision making support was provided?*

- ☐ Support data coordinator
- ☐ Support Leadership team in using data
- ☐ Support Internal Coach in using data and data analysis
- ☐ Support classroom/student level staff in using data and data analysis
- ☐ Other - Write In (Required): _____ *

Logic: Hidden unless: Question "What types of support were provided? (check all that apply)" #6 is one of the following answers ("Behavior Systems")
10) What type of behavior systems support was provided?*

- ☐ Support behavior specialist
- ☐ Support Leadership Team in design and implementation of systems for behavior support
- ☐ Other - Write In (Required): _____ *

Logic: Show/hide trigger exists. Hidden unless: Question "What types of support were provided? (check all that apply)" #6 is one of the following answers ("Family Group Training")
11) What type of family group training was provided?*

- ☐ Provide information to schools/districts on family training
- ☐ Provide Positive Solutions training independently
- ☐ Co-present Positive Solutions with District personnel
- ☐ Other - Write In (Required): _____ *

Logic: Hidden unless: Question "What type of family group training was provided?" #11 is one of the following answers ("Provide Positive Solutions training independently")
12) Which modules were covered in the training?*

- ☐ Module 1
- ☐ Module 2

- ☐ Module 3
- ☐ Module 4
- ☐ Module 5
- ☐ Module 6

Validation: Must be numeric
Logic: Hidden unless: Question "What type of family group training was provided?" #11 is one of the following answers ("Provide Positive Solutions training independently")
13) How many people participated in the Positive Solutions training?*

Description of the Activities

Page exit logic: Page Logic**IF:** Question "Have you been identified as an external coach to provide supports to this district/program?" #18 is one of the following answers ("No") **THEN:** Jump to page 4 - Thank You! Flag response as complete

14) Description of the activities (If a meeting took place as part of this contact please email an agenda with an updated status for each agenda item to your ESE contact (Martha Daigle or Sarah Geldart): *

Validation: %s format expected
15) Prior approval submitted*

16) List the district/school staff names and roles of individuals who participated in this contact. If this was a family training, provide the organization(s) who participated. *

17) Are there any additional next steps or notes you would like to share with ESE as a result of this contact?

Please email any documentation of your activities such as a presentation, meeting agenda, or list of next steps to Sarah Geldart at sarah.geldart@doe.mass.edu and Martha Daigle at mdaigle@doe.mass.edu upon completion of this contact record.

18) Have you been identified as an external coach to provide supports to this district/program?*

- ○ Yes
- ○ No

Current Implementation Status

To the best of your knowledge, please describe the identified district's current Preschool PBS through Pyramid Strategies implementation status.

Validation: Must be numeric

19) Number of schools/programs participating in this district:

20) Number of classrooms with trained teachers fully implementing (provide details):

21) Number of classrooms with trained teachers who have adopted a portion of the model or materials (provide details):

22) Number of classrooms with teachers who have not had formal training and have adopted a portion of the model or materials (provide details):

23) Describe any recent school or program-wide PBS activities:

24) Have any new PBS products been created for the classroom, school or program? (If so, please describe them):

25) Do you have any additional information about the use of the TPOT in the district? Have coaches used the tool and if so, do you know the results?

26) What types of data does the Leadership Team utilize for planning?

- Indicator 7 (Preschool Outcomes)
- Behavior Incident Reports
- IEP Data
- Behavior Intervention Plan Data
- Removal Data
- Other - Write In (Required): *

27) How often does the Leadership Team review data for planning?

- At every Leadership Team meeting
- Occasionally
- Rarely
- It is a built in part of their system
- Other - Write In (Required): *

28) How does the Leadership Team use their data?

29) As an external coach, how are you supporting the Leadership Team's data use?

Thank You!

Thank you completing this contact record. Please contact Sarah Geldart at sarah.geldart@doe.mass.edu with any questions.

Johns Hopkins University
Homewood Institutional Review Board (HIRB)

Informed Consent Form.

Title:	Understanding the Implementation of the Pyramid Model in Massachusetts
Principal Investigator:	Sarah Geldart, Doctor of Education Candidate, Johns Hopkins University
Date:	July 1, 2016

PURPOSE OF RESEARCH STUDY:

The purpose of this research study is to understand the implementation of the Pyramid Model for Supporting Social Emotional Competence in Infants and Young Children in Massachusetts during the 2015-2016 school year.

We anticipate that approximately 1-2 people will participate in this study, including affiliates from the Pyramid Model Consortium.

PROCEDURES:

This study will involve face-to-face or telephone interviews of national staff.

Time Required: approximately 25-45 minutes.

RISKS/DISCOMFORTS:

There are no anticipated risks to interviewees.

BENEFITS:

There are no direct benefits to you from participating in this study. However, this study may benefit young children with disabilities in Massachusetts if the results lead to a better understanding of implementation with fidelity of evidenced based practices that can be put in place to support the social emotional development of these students and how programs that support this population can be expanded.

VOLUNTARY PARTICIPATION AND RIGHT TO WITHDRAW:

Your participation in this study is entirely voluntary: You choose whether to participate. If you decide not to participate, there are no penalties, and you will not lose any benefits to which you would otherwise be entitled.

If you choose to participate in the study, you can stop your participation at any time, without any penalty or loss of benefits. If you want to withdraw from the study, please let the interviewer know at any time during the interview.

CONFIDENTIALITY:

Any study records that identify you will be kept confidential to the extent possible by

law. The records from your participation may be reviewed by people responsible for making sure that research is done properly, including members of the Johns Hopkins University Homewood Institutional Review Board and officials from government agencies such as the National Institutes of Health and the Office for Human Research Protections. (All of these people are required to keep your identity confidential.) Otherwise, records that identify you will be available only to people working on the study, unless you give permission for other people to see the records. Participants names and roles will not be provided in any reports of the interviews or the study in its entirety.

COMPENSATION:

You will not receive any payment or other compensation for participating in this study.

IF YOU HAVE QUESTIONS OR CONCERNS:

You can ask questions about this research study now or at any time during the study, by talking to the researcher working with you or by calling Sarah Geldart at 617-710-0210.

If you have questions about your rights as a research participant or feel that you have not been treated fairly, please call the Homewood Institutional Review Board at Johns Hopkins University at (410) 516-6580.

SIGNATURES

WHAT YOUR SIGNATURE MEANS:

Your signature below means that you understand the information in this consent form. Your signature also means that you agree to participate in the study.
By signing this consent form, you have not waived any legal rights you otherwise would have as a participant in a research study.

Participant's Signature **Date**

Signature of Person Obtaining Consent (Investigator or HIRB Approved Designee) **Date**

Interview Protocol: Understanding the Statewide Implementation of the Pyramid Model in Massachusetts

Project: Understanding the implementation of the Pyramid Model to address poor social emotional outcomes for children with disabilities aged 3-5 in Massachusetts.

Interviewee (Title and Name): _____

Interviewer: _____

Survey Section(s) Used:

_____ A: Interview Background

_____ B: Institutional Perspective

_____ C: Existing Systems and Supports

_____ D: Concluding Thoughts

Other Topics Discussed:_____

_____

Documents Obtained: _____

_____

_____

Post Interview Comments or Leads:

_____

NOTE: Expand space between questions and prompts as necessary to capture notes from the interview.

Early Childhood Social Emotional Outcomes in Massachusetts

Introductory Protocol Script:

This interview is intended to explore your understanding and thoughts around the implementation of the Pyramid Model in Massachusetts at the statewide level. In order to participate in this interview I will need you to read and sign this letter of consent. This document basically states that: (a) any information you provide will be kept confidential, (b)

your participation is completely voluntary and you may end the interview at any time, and (c) no harm to you is anticipated based on your participation. Thank you for agreeing to participate.

This interview should last no longer than 45 minutes.

Introduction

You have been identified as a key informant who has a great deal of information to share about the process and fidelity of implementation of the Pyramid Model in the Massachusetts public school context. This research project seeks to understand your perspective on how the state approached implementation of this model during the initial year of implementation (2015-2016). It also seeks to develop a contextualized understanding of differences in Massachusetts' implementation compared to other, similar states and how the model was intended to be implemented. This project does not intend to evaluate the performance of the Massachusetts Department of Elementary and Secondary Education or school districts, but seeks to understand your perspective on this implementation and how the state utilized resources from the Consortium, CSEFEL, and TACSEI.

A. Interviewee Background

1. What is your current relationship to the Pyramid Model Consortium?

2. Can you describe your experience with the implementation of the Pyramid Model?

Probes: Do you have previous experience working with other states implementing this
model?

As a state employee, how familiar do you feel with the current challenges and issues faced at the local level?

B. Institutional Perspective

1. What is your current understanding of the statewide implementation of the Pyramid Model in Massachusetts?

Probes: What is your sense of how the model is being implemented at the state level?

Do you feel as though it is being implemented/used as intended?

2. In what ways has MA ESE supported the use of the Pyramid Model at the district level?

Probes: What professional development and data quality initiatives have the state supported?

In your opinion are there any additional state-level supports necessary for districts to use this process effectively?

3. MA ESE has chosen to implement the Pyramid Model as part of its State Systemic Improvement Plan or SSIP. What is your understanding of how Massachusetts' implementation aligns with the research literature on the implementation of this practice?

Probes: In what ways does the implementation align with the literature? In what ways is in not aligned?

Are the components of the intervention being delivered as designed?

3. Is the implementation of the Pyramid Model and its associated trainings in this state aligned with how the model was intended to be delivered?

Probe: If not, how is it different?

4. The quality of delivery at statewide trainings is very important for ensuring implementation fidelity. How does the Pyramid Model Consortium support high-quality professional development and technical assistance in the state?

Probes: Are the intended methods and resources being used to deliver the intervention consistent with how the intervention was designed?

What training and backgrounds do the national trainers bring to their work in Massachusetts?

C. Existing Systems and Supports

1. What activities have been undertaken by MA ESE in the past year to support the high-fidelity implementation of the Pyramid Model in Massachusetts at the state level?

Probes: Do you think these activities were sufficient to support high-quality implementation?

Do these activities align with your recommended practices? If not, how do they differ?

2. What do you believe are the key state-level initiatives to support social emotional outcomes for young children with disabilities?

Probes: How can these be strengthened or improved?

Are there any areas where additional state-level activities may need to be introduced?

4. How do the existing systems to support social emotional outcomes for this population in Massachusetts compare to other states?

Probes: Are you aware of any states that are doing exceptionally well in supporting these outcomes?

Are there any national initiatives that could impact the work in Massachusetts or be leveraged to support improved social emotional outcomes in this population?

D. Concluding Thoughts

1. What is your overall impression of the implementation of this model in Massachusetts in the most recent school year?

2. What are the next steps for MA ESE in supporting improved social emotional outcomes for young children with disabilities through the Pyramid Model?

Probes: How can the mixed-delivery system be leveraged to support this improvement?

What role will classroom professionals including teachers, service providers, and paraprofessionals play in this work?

2. Is there any additional information you would like to provide about the work in Massachusetts to improve social emotional outcomes for children with disabilities aged 3-5?

Conclusion Protocol Script:

Thank you for your participation in this interview. Your feedback will be very helpful in understanding the current status and activities underway in Massachusetts to support improved social emotional outcomes for children with disabilities aged 3-5. If you have anything you would like to add or additional questions after this interview please do not hesitate to contact me.

Post Interview Comments and/or Observations:

Curriculum Vitae

SARAH WHITMAN GELDART
16 Rice St. #2, Brookline, MA 02445 * 617-710-0210 * sarah.whitman@gmail.com

EDUCATION

Ed.D., Anticipated 2017, Johns Hopkins University
Ed.M., Education Policy and Management: 2012, Harvard University
M.A., Religion and Society: 2007, Boston University
B.A., Double Major in Psychology and Religion: 2005, Boston University

PROFESSIONAL EXPERIENCE

Massachusetts Department of Elementary and Secondary Education, Special Education Planning and Policy Development Office, Malden, MA (October 2012 – March 2017)
Early Childhood Outcomes Coordinator and Data Analyst (Educational Specialist C)

- Project Co-Lead on the State Systemic Improvement Plan, a federally-mandated multi-year initiative to improve outcomes for children with disabilities. Responsibilities include data analyses, stakeholder engagement, authoring federal reports, and development and implementation of project evaluation plan
- Responsible for the statewide collection, analysis, and federal reporting for three indicators of state level performance in serving children with disabilities: Indicators 6, 7, and 17
- Manages a statewide cadre of eight trainers including hiring, onboarding and ongoing supervision
- Oversees the statewide implementation of the Pyramid Model, a tiered system of support, including the procurement of national trainers, recruitment of districts to participate, and training of state level coaches
- Created and supports the Early Childhood Special Education Systemic Program Improvement Grant, a muti-year grant program that has disbursed over $4 million to districts in the state
- Collaborated with contractors to create the first special education results report for the state's longitudinal data analysis and reporting tool, EdWin Analytics
- Developed a multi-year statewide professional development grant and related training opportunities for over 400 school districts to support performance improvement and transformational practices
- Develops inter-agency statewide initiatives to support improved longitudinal outcomes for students
- Designed more than twenty different face-to face and online professional development trainings to increase statewide capacity of educators, including courses on data use for continuous improvement
- Provides in-office training to further develop staff skills in collecting, analyzing, and interpreting special education data, including assessing whether implementation of an initiative is successful
- Completed a statewide, one year Program Director training to further develop skills in program management and strategies for maximizing organizational impact

Program on Negotiation (PON) at Harvard Law School, Cambridge, MA (July 2007-July 2012)
Research Associate and Student Interest Group (SIG) Program Coordinator
Promoted from Program Assistant

- Worked with the 8 major PON research initiatives to track progress and ongoing research
- Initiated, implemented, and evaluated student-oriented activities including the PON Career Series, Internships, Student Discussion Groups, grants, and the PON Fellowship Programs
- Facilitated student involvement in PON which included over 1200 Student Interest Group members and served as a resource to students interested in negotiation and dispute resolution
- Oversaw all SIG financials including budgeting, disbursements, and university reporting
- Supervised graduate student assistants, volunteers, and interns
- Outreach responsibilities included liaising with partner academic institutions, designing and disseminating relevant materials, and promoting relationships among faculty, mentors and students

Facing History and Ourselves, Brookline, MA (September 2006-July 2007)
Research Coordinator for Evaluation
Promoted from Research Assistant for Evaluation

- Assisted in development, recruitment, and implementation of national experimental study with over 100 school and 250 teacher participants
- Advised nine regional directors regularly regarding ongoing evaluation research, including development of new organizational policies and procedures
- Collected, analyzed, and generated regular reports on evaluation surveys
- Assisted in developing and writing presentations for internal and external audiences
- Implemented new departmental organizational system including an online database
- Generated presentations and reports for directors, the organizational board, and the public

The Center for the Study of Psychology and Religion at the Danielsen Institute, Boston, MA (September 2004-July 2005, September 2006-June 2007)
Graduate Research Assistant
Promoted from Undergraduate Research Assistant

- Supported faculty qualitative and quantitative research
- Conducted interviews of experiment participants and community members for research projects
- Assisted in the development of grant proposals and edited documents and webpage
- Prepared presentations for major national conferences

TEACHING EXPERIENCE

Johns Hopkins University, Baltimore, MD (August 2015 – January 2016)

Teaching Assistant

- Serves as a teaching assistant in the online doctoral course, Disciplinary Approaches to Education
- Responsibilities include managing online discussion forums, grading, hosting virtual course sessions on course content and writing skills, and one-on-one coaching to students

CONFERENCE PRESENTATIONS

- Invited Presentation: Implementing and Scaling the Pyramid Model in Massachusetts' Public Schools - from Policy to Practice. *Massachusetts Pyramid Model Partnership Summit.* Westford, MA, April 2016
- Panel Presentation: Digging Deeper: Helping Programs Use Child Outcomes Data to Improve Services, *Improving Data, Improving Outcomes Early Childhood National Conference,* New Orleans, LA, September 2014
- Panel Presentation: Special Education Policy in Texas and Massachusetts. *Harvard Graduate School of Education Student Research Conference.* Cambridge, MA, April 2012

GOVERNMENT APPOINTMENTS

Commission for the Disabled, Brookline, MA (January 2012-January 2017)

- As a member of the commission, works with the Massachusetts Office on Disability to carry out programs and activities designed to integrate people with disabilities into the community

VOLUNTEER EXPERIENCE

Special Education Surrogate Parent Program, Massachusetts Department of Elementary and Secondary Education (May 2011-October 2012)
Volunteer Special Education Surrogate Parent

Harvard Mediation Program, Cambridge, MA (January 2010-October 2012)
Community Mediator and Trainer